

# Movie Raitings

---

ELISEO ORELLAN ANGUIANO

# Movie Raitings

---

1. Base de Datos de GroupLens
2. Bases de Datos
  1. Raitings → 58K
  2. Personalities → 1.8K
3. Problema a Resolver

Optimizar el catalogo de películas mediante conocer cuales son los perfiles más populares y cuales son sus preferencias de películas.

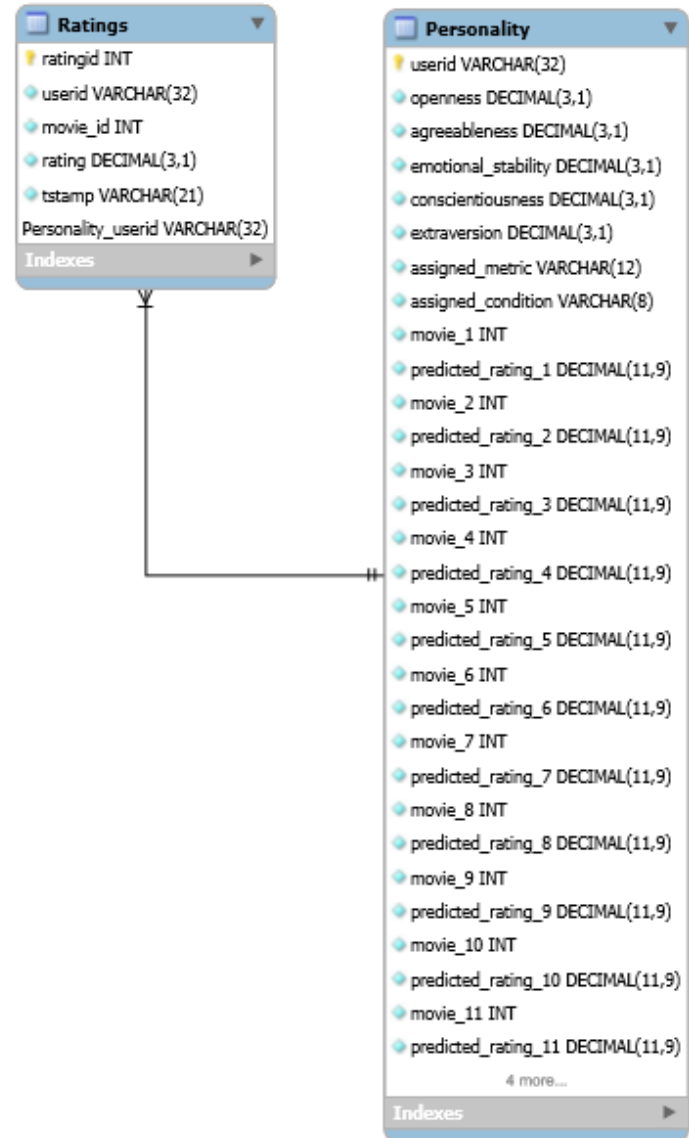
# Estructura de Bases de Datos

## 1. Ratings

1. Usuarios
2. Películas
3. Valoración (Rating)

## 2. Personalidad

1. Dimensión Personalidad
2. Métrica de Personalidad
3. Nivel de Personalidad
4. Película
5. Predicción de Rating



# Registros Muestra

MySQL Workbench

MyFirstConnection\_Eliseo x MySQL Model\* x EER Diagram x

File Edit View Query Database Server Tools Scripting Help

Navigator: PERSONALITY Postwork Session 1 - Eliseo Orell...

1 • `SELECT * FROM MovieRatings.Personality;`

Result Grid

| userid                            | openness | agreeableness | emotional_stability | conscientiousness | extraversion | assigned_metric | assigned_condition | movie_1 | predicted_rating_1 | movie_2 | predicted_rating_2 |
|-----------------------------------|----------|---------------|---------------------|-------------------|--------------|-----------------|--------------------|---------|--------------------|---------|--------------------|
| 005fe8678214011d7f92e51f9a546d40  | 5.5      | 6.0           | 2.0                 | 4.5               | 1.0          | popularity      | low                | 94466   | 4.172660863        | 77658   | 4.199668475        |
| 0066fac81b62656f032c085d96e378f4  | 4.5      | 4.5           | 2.5                 | 4.5               | 4.0          | diversity       | low                | 69844   | 4.719698785        | 54001   | 4.711537968        |
| 00fa91e202f5e48aa34c05d97867fa74  | 6.5      | 3.5           | 3.0                 | 5.0               | 5.0          | serendipity     | medium             | 1023    | 4.070440702        | 38038   | 3.904671131        |
| 011aedbea90fb3b6d1e7a47526b3bee6  | 5.5      | 4.0           | 3.0                 | 1.5               | 2.5          | serendipity     | high               | 32840   | 3.980606322        | 112940  | 3.676377893        |
| 013325441cd1755b18c58b8007646bb1  | 6.5      | 6.5           | 6.0                 | 7.0               | 5.0          | popularity      | low                | 112852  | 3.946324100        | 111759  | 3.897057460        |
| 0160f8c5607073505cb7be7157136a1b  | 5.0      | 3.0           | 5.5                 | 4.5               | 4.5          | serendipity     | high               | 129514  | 4.499130758        | 113741  | 4.424165064        |
| 01a88049c6c3c11cf94e1921a042a2f4  | 6.0      | 3.0           | 6.0                 | 6.5               | 3.0          | diversity       | medium             | 64034   | 4.444692782        | 1077    | 4.460173875        |
| 01aedde55cfaf3825b8f1e162fda5225  | 5.5      | 4.0           | 3.0                 | 3.0               | 1.5          | popularity      | medium             | 44555   | 4.130520993        | 27773   | 3.981667705        |
| 01e4e18e2e21b84ee1c5813fd5d3797f  | 6.5      | 4.5           | 3.0                 | 4.5               | 5.5          | popularity      | medium             | 45517   | 4.029183231        | 59784   | 4.131958185        |
| 02077bd159c08095b9dc967cab60829e  | 6.0      | 5.5           | 4.5                 | 5.5               | 2.5          | popularity      | low                | 108583  | 3.916236374        | 108979  | 3.974787333        |
| 022047320a00d607009323875a19face  | 5.0      | 6.0           | 6.0                 | 5.0               | 4.5          | all             | default            | 1203    | 4.373463621        | 923     | 4.645019767        |
| 022b7d7h0584hhf5e4c7d85ed60h3fa16 | 5.5      | 4.0           | 5.0                 | 3.5               | 4.0          | all             | default            | 318     | 4.953904055        | 58559   | 4.821731369        |

# Registros Muestra

MySQL Workbench

MyFirstConnection\_Eliseo x MySQL Model\* x EER Diagram x

File Edit View Query Database Server Tools Scripting Help

Navigator: SCHEMAS

Filter objects

- JDA
- JDA\_world
- JV
- MET
- Modulos2
- MovieRatings
  - Tables
    - Personality
    - Ratings
  - Views
  - Stored Procedures
  - Functions
- PRUEBA
- PRUEBA1
- prueba2
- sakila
- Seguimiento\_Requets
- SV
- SVSchema
- sys
- tienda

Administration Schemas

Information

Table: Ratings

Columns:

- ratingid int PK
- userid varchar(32)
- movie\_id int
- rating decimal(3,1)
- tstamp varchar(21)

Object Info Session

Personality Ratings x Postwork Session 1 - Eliseo Orell...

1 • SELECT \* FROM MovieRatings.Ratings;

Result Grid

|    | ratingid                        | userid | movie_id | rating              | tstamp |
|----|---------------------------------|--------|----------|---------------------|--------|
| 1  | 8e7ceb9a234c064b75016249f2ac65e | 1      | 5.0      | 2001-09-10 17:19:56 |        |
| 2  | 8e7ceb9a234c064b75016249f2ac65e | 2      | 4.0      | 2001-09-28 11:34:55 |        |
| 3  | 8e7ceb9a234c064b75016249f2ac65e | 3      | 4.0      | 2001-09-28 11:42:50 |        |
| 4  | 8e7ceb9a234c064b75016249f2ac65e | 5      | 5.0      | 2001-09-28 11:27:30 |        |
| 5  | 8e7ceb9a234c064b75016249f2ac65e | 6      | 4.0      | 2002-01-07 18:12:02 |        |
| 6  | 8e7ceb9a234c064b75016249f2ac65e | 7      | 4.0      | 2001-09-21 14:09:24 |        |
| 7  | 8e7ceb9a234c064b75016249f2ac65e | 10     | 4.0      | 2001-09-29 18:13:15 |        |
| 8  | 8e7ceb9a234c064b75016249f2ac65e | 11     | 4.0      | 2001-09-21 11:54:21 |        |
| 9  | 8e7ceb9a234c064b75016249f2ac65e | 12     | 4.0      | 2001-09-29 18:58:15 |        |
| 10 | 8e7ceb9a234c064b75016249f2ac65e | 13     | 4.0      | 2001-09-14 15:29:46 |        |
| 11 | 8e7ceb9a234c064b75016249f2ac65e | 14     | 3.5      | 2004-05-10 22:38:37 |        |
| 12 | 8e7ceb9a234c064b75016249f2ac65e | 15     | 2.0      | 2001-09-29 19:02:38 |        |
| 13 | 8e7ceb9a234c064b75016249f2ac65e | 16     | 4.0      | 2001-09-19 09:22:33 |        |

Output

Action Output

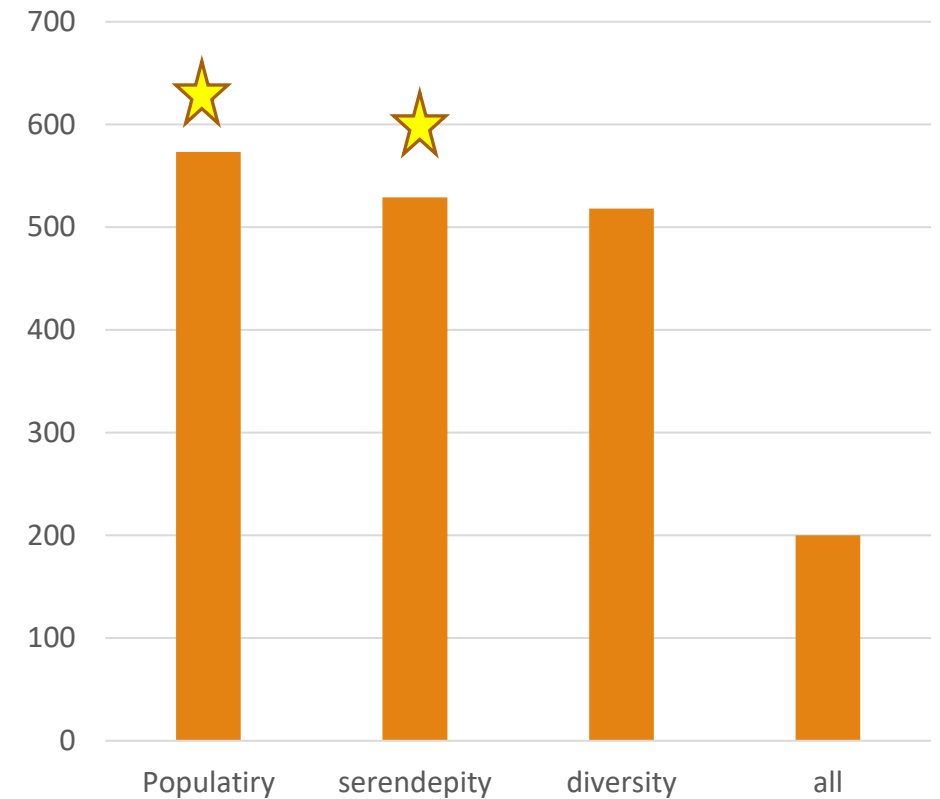
| # | Time     | Action  | Message               | Duration / Fetch       |
|---|----------|---|-----------------------|------------------------|
| 2 | 13:30:10 | EXPLAIN SELECT * FROM MovieRatings.Personality                              | OK                    | 0.000 sec              |
| 3 | 13:30:10 | EXPLAIN FORMAT=JSON SELECT * FROM MovieRatings.Personality                  | OK                    | 0.000 sec              |
| 4 | 13:32:44 | ANALYZE TABLE 'MovieRatings'.Personality                                    | OK                    | 0.000 sec              |
| 5 | 13:40:29 | SELECT assigned_metric , count(*) FROM Personality group by assigned_metric | 4 row(s) returned     | 0.078 sec / 0.000 sec  |
| 6 | 13:40:50 | SELECT * FROM MovieRatings.Personality                                      | 1820 row(s) returned  | 0.156 sec / 3.391 sec  |
| 7 | 13:42:25 | SELECT * FROM MovieRatings.Personality                                      | 1820 row(s) returned  | 0.297 sec / 3.797 sec  |
| 8 | 13:42:30 | SELECT * FROM MovieRatings.Ratings  | 58938 row(s) returned | 0.156 sec / 32.313 ... |

# ¿Cuál es el perfil más popular?

```
9      -- cuales son las personalidades mas comunes?
10
11 •    SELECT assigned_metric , count(*)
12      FROM Personality
13      group by assigned_metric;
14
```

< Result Grid | Filter Rows: | Export: | Wrap Cell C

|   | assigned_metric | count(*) |
|---|-----------------|----------|
| ▶ | popularity      | 573      |
|   | diversity       | 518      |
|   | serendipity     | 529      |
|   | all             | 200      |



# ¿Cual es perfil con mayor nivel de “openness”?

---

```
15 -- cual es perfil con mas openness?
16 • SELECT assigned_metric , avg(openness) as apertura
17 FROM Personality
18 group by assigned_metric
19 ORDER BY apertura DESC;
```

| Result Grid |                 | Filter Rows: | Export: | Wrap Cell Content |
|-------------|-----------------|--------------|---------|-------------------|
|             | assigned_metric | apertura     |         |                   |
| ▶           | all             | 5.44250      |         |                   |
|             | popularity      | 5.40227      |         | ★                 |
|             | diversity       | 5.35425      |         |                   |
|             | serendipity     | 5.33270      |         | ★                 |

# ¿Cuál perfil tiene mayor nivel de “agreeableness”?

Personality Ratings Postwork Sesion 1 - Eliseo Orell...

-- cual es perfil con mas agreeableness?

```
21 • SELECT assigned_metric , avg(agreeableness) as agradable
22 FROM Personality
23 group by assigned_metric
24 ORDER BY agradable DESC;
```

Result Grid Filter Rows: Export: Wrap Cell Content:

|   | assigned_metric | agradable |
|---|-----------------|-----------|
| ▶ | diversity       | 4.28475   |
|   | popularity      | 4.27923   |
|   | all             | 4.21750   |
|   | serendipity     | 4.07750   |

★

★



# ¿Cuál es el perfil con mayor nivel de “Emotional Stability”

Personality Ratings Postwork Sesion 1 - Eliseo Orell...

-- cual es perfil con mas emotional stability?

```
26 • SELECT assigned_metric , avg(emotional_stability) as Estabilidad
27 FROM Personality
28 group by assigned_metric
29 ORDER BY Estabilidad DESC;
```

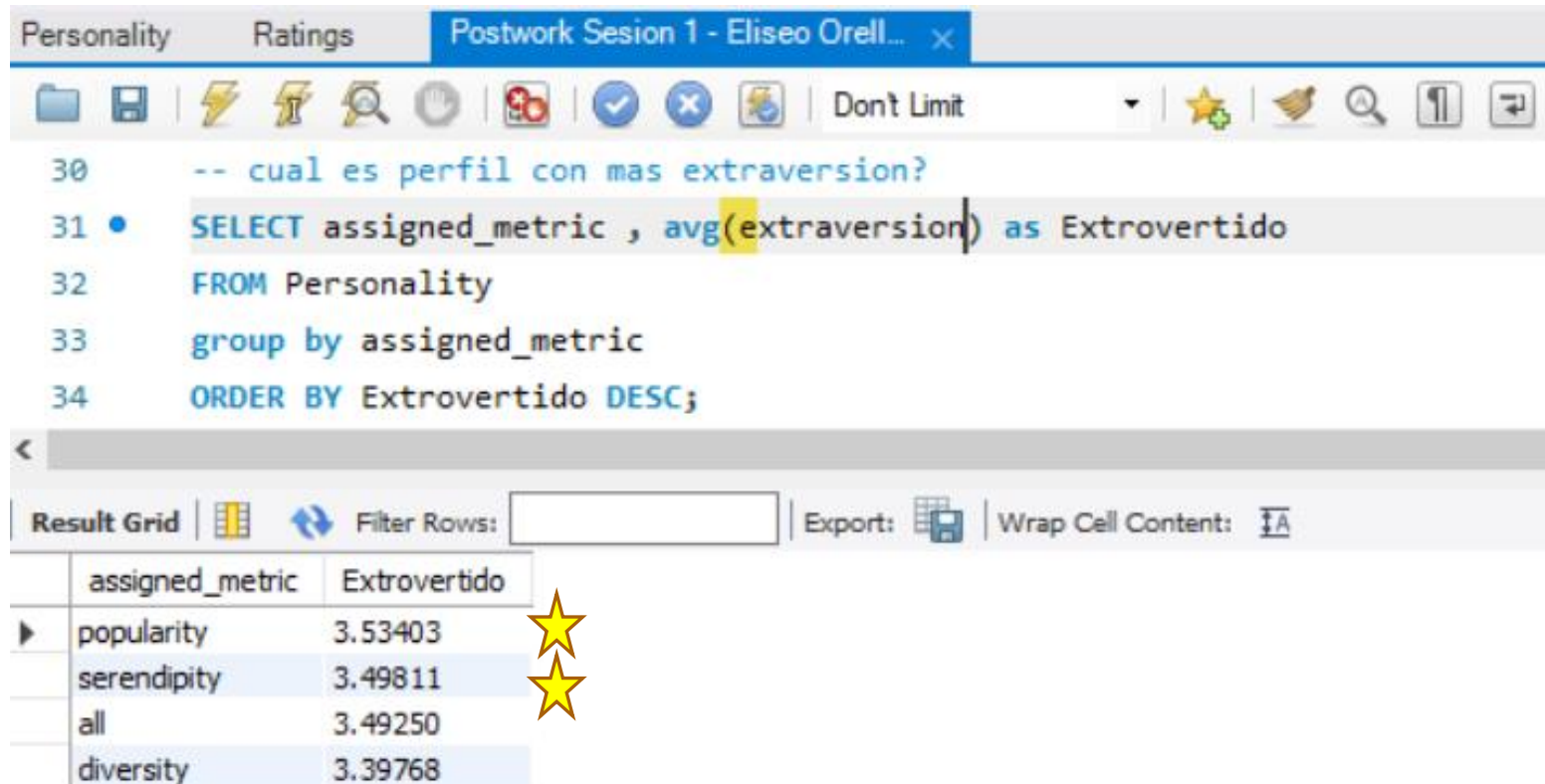
Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

|   | assigned_metric | Estabilidad |
|---|-----------------|-------------|
| ▶ | diversity       | 4.62838     |
|   | all             | 4.58250     |
|   | popularity      | 4.53490     |
|   | serendipity     | 4.52552     |

★

★

# ¿Cual es perfil con mas extraversión?



The screenshot shows a database query tool interface. The top bar has tabs for 'Personality', 'Ratings', and 'Postwork Sesion 1 - Eliseo Orell...'. Below the tabs is a toolbar with various icons. The main area displays a SQL query:

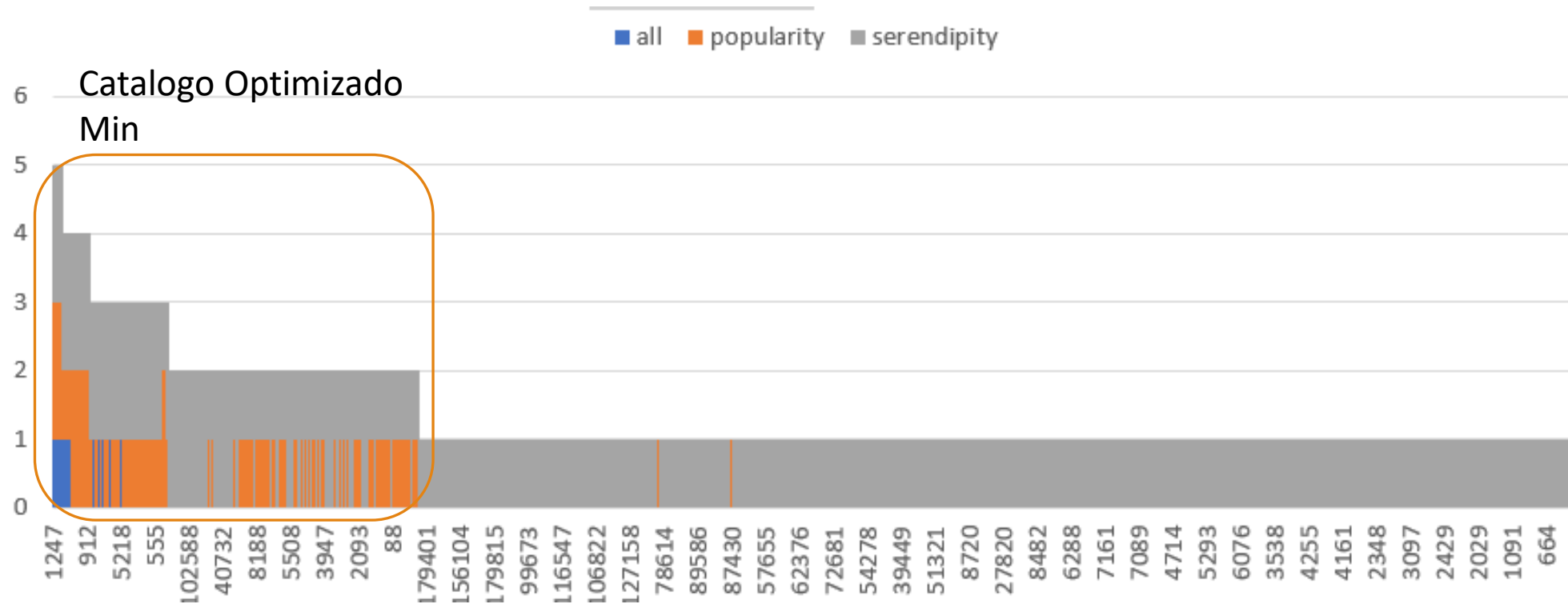
```
30  -- cual es perfil con mas extraversión?  
31  • SELECT assigned_metric , avg(extraversión) as Extrovertido  
32  FROM Personality  
33  group by assigned_metric  
34  ORDER BY Extrovertido DESC;
```

Below the query is a 'Result Grid' section. It includes a 'Filter Rows' input field, an 'Export' button, and a 'Wrap Cell Content' checkbox. The results are displayed in a table:

|   | assigned_metric | Extrovertido |
|---|-----------------|--------------|
| ▶ | popularity      | 3.53403      |
|   | serendipity     | 3.49811      |
|   | all             | 3.49250      |
|   | diversity       | 3.39768      |

Two yellow stars are placed next to the 'popularity' and 'serendipity' rows, indicating they are the top two profiles with the highest extraversion scores.

# Optimizar el catalogo de películas para que pueda ser del agrado a los perfiles de acuerdo a la preferencia de los usuarios



# Conclusiones

---

1. Optimizar catalogo de películas por el grado de preferencia que tiene y buscar incluir más películas que encajen con un perfil alto de “Extravesión”
2. Personalmente se me facilita más el uso de SQL por la sintaxis y la estructura visual
3. El performance de Mongo es impresionante y me gustaría profundizar mi conocimiento
4. El mayor trabajo empieza en definir bien las preguntas y el porque de su importancia