Proyecto: QS Rankings en el mundo y Mexico

1. Definición del proyecto (Fuentes y problemas a resolver)

1) <u>Fuentes de BD:</u>

Los datos provienen de un dataset de Kaggle con información del ranking anual QS World Universities: https://www.kaggle.com/prasertk/qs-world-university-rankings-

2021?select=QS+World+University+Rankings+2017.csv

2) Problema a resolver:

La información proviene de una investigación realizada por la Secretaria de Educación que busca identificar a las universidades mejor calificadas en el mundo, compararlas con las mejores Universidades de México así como obtener una idea del perfil que da el sistema educativo frente al mundo para así comenzar a fijar metas realistas para implementar mejoras al sistema y preparar un programa de becas para ayudar a alumnos ejemplares a obtener oportunidades en dichas universidades.

2. Definición de la base de datos

La data proviene de 3 archivos:

- a) QS Ranking 2021: El archivo QS Ranking 2021.csv contiene toda la información relevante sobre la clasificación de las universidades en el QS World University 2021.
- b) QS Ranking 2020: El archivo QS Ranking 2020.csv contiene toda la información relevante sobre la clasificación de las universidades en el QS World University 2020.

Diagrama Entidad Relación

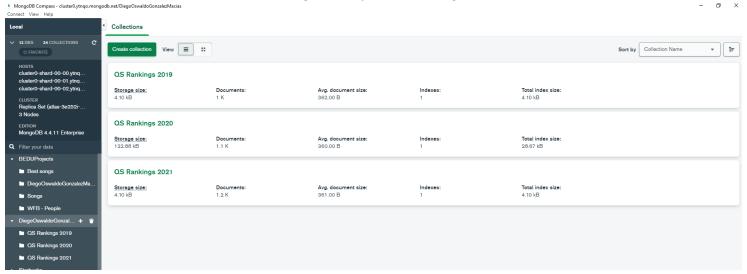
Glosario

QS Ranking 2020 y QS Rankings 2021

- university (String): Nombre de la universidad.
- year (Decimal): Año en que se realizó la clasificación.
- rank_display (Decimal): posición en la clasificación que se encuentra basada en su calificación. Puede haber empates.
- university (String): Nombre de la universidad.
- **score** (Decimal): Las instituciones son evaluadas de acuerdo con las siguientes seis métricas:
- Reputación académica (40%)
- Reputación del empleador (10%)
- Ratio profesorado/alumnos (20%)
- Citas por profesorado (20%)
- Ratio de profesorado internacional (5%)
- Ratio de estudiantes internacionales (5\$)
 - **link** (String): Hipervínculo a la Home Page de la universidad en el sitio web de QS.
 - **country** (Decimal): El pais de la institución.
 - city (Decimal): La ciudad de la institución.
 - **region** (Decimal): Región donde se encuentra la universidad (puede no aplicar).
 - **logo** (Decimal): Hipervínculo al logo de la institución en la página web de QS Top Universities.

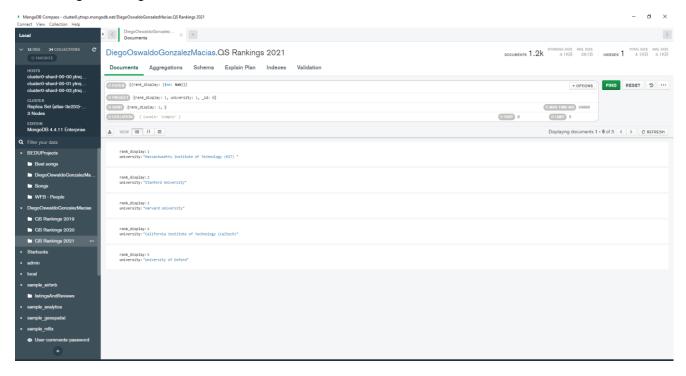
Creación de BBDD en Mongo DB

Colecciones de QS Rankings 2020 y QS Ranking 2021

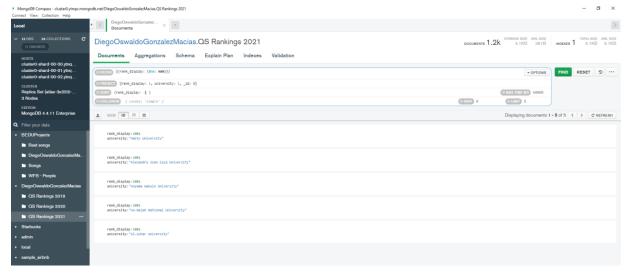


3. Consultas

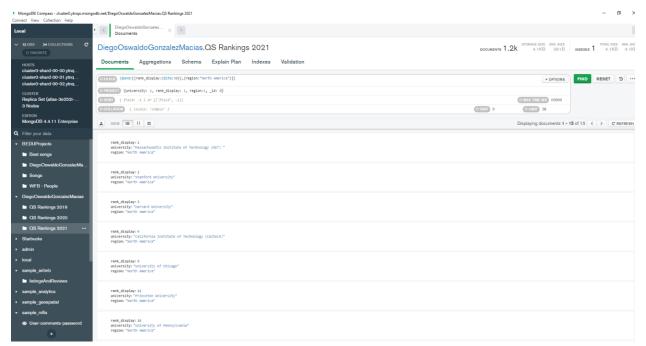
1) ¿Cuáles han sido las 5 primeras universidades en 2021, excluyendo aquellas que no tienen calificación o clasificación?



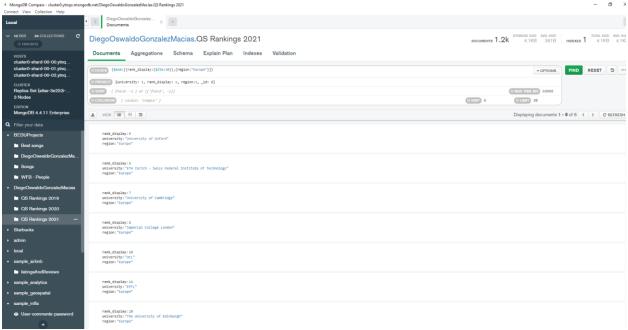
2) ¿Cuáles han sido las 5 universidades que han quedado al final de la clasificación en 2021?



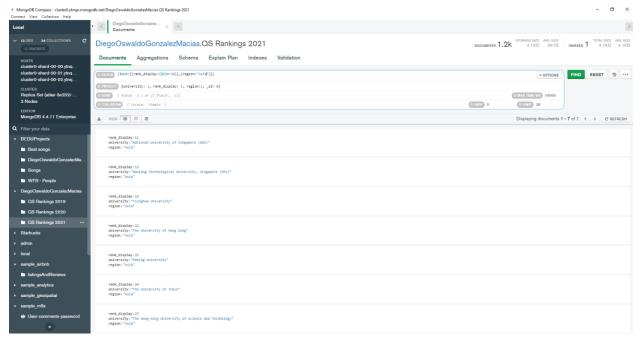
3) ¿Cuántas escuelas en el top 30 son de la región de Norteamérica y que clasificación tienen?



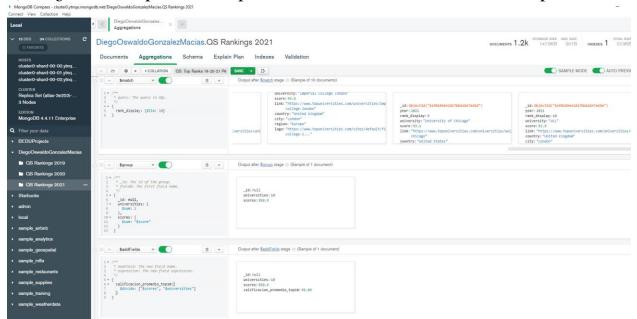
4) ¿Cuántas escuelas en el top 30 son de la región de Europa y que clasificación tienen?



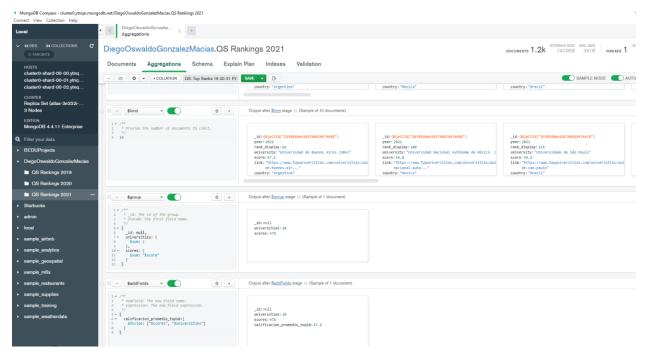
5) ¿Cuántas escuelas en el top 30 son de la región Asia y que clasificación tienen?



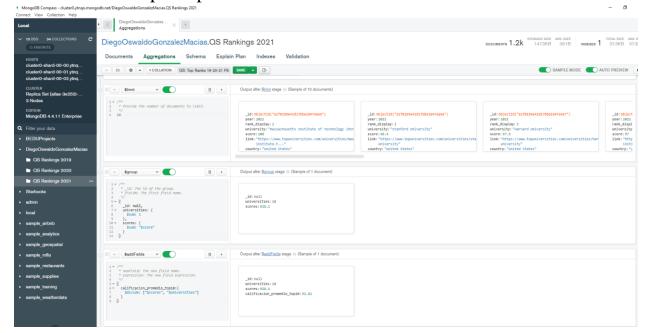
6) ¿Cuál es la puntuación promedio de las universidades en el top 10?



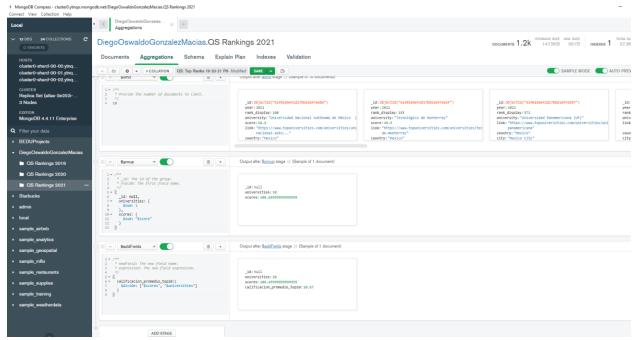
7) ¿Cuál es la puntuación promedio de las 10 mejores universidades Latinoamericanas?



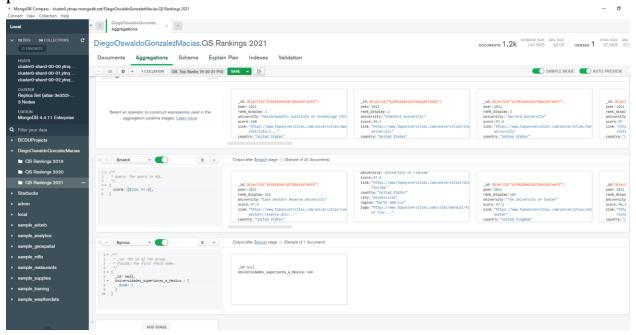
8) ¿Cuál es la puntuación promedio de las 10 universidades norteamericanas principales?



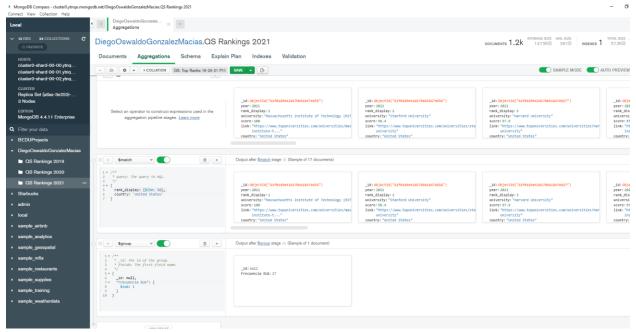
9) ¿Cuál es el promedio de las 10 mejores Universidades mexicanas?



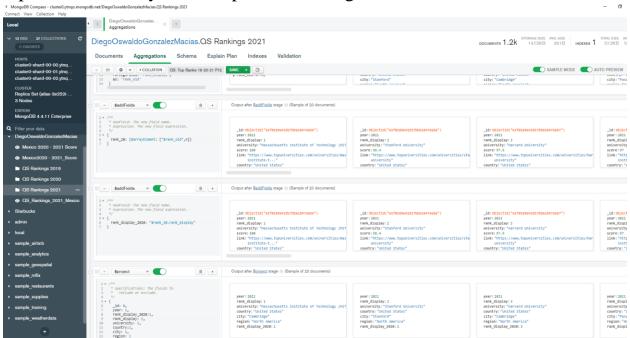
10) ¿Cuáles son las universidades con una puntuación superior al promedio latinoamericano?



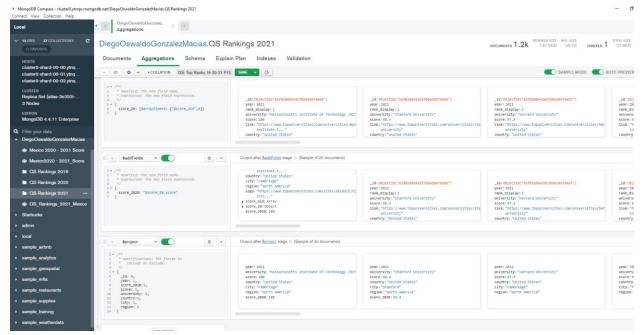
11) ¿Cuántas universidades del top 50 son de Estados Unidos?



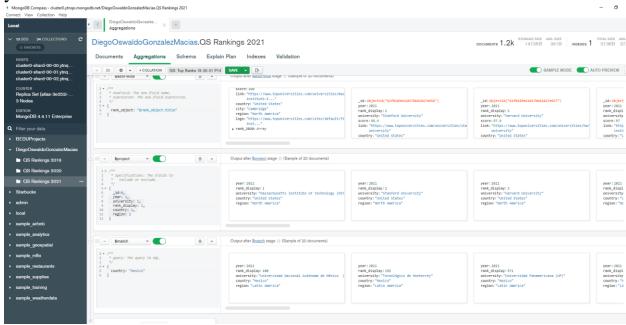
12) ¿Cuál ha sido la diferencia o cambio entre las universidades en el top 10 entre 2020 y 2021 respecto a ranking?



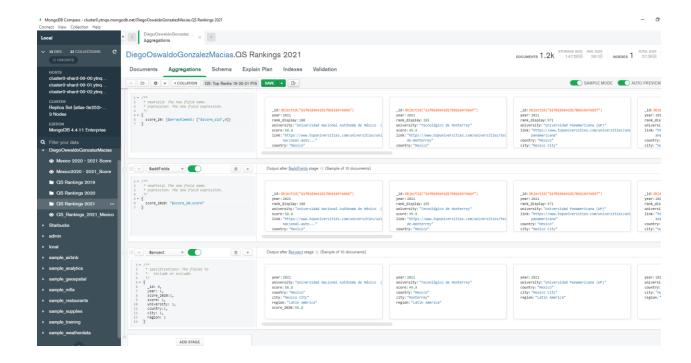
13) ¿Cuál ha sido la diferencie en calificación entre las universidades en el top 10 entre 2020 y 2021?



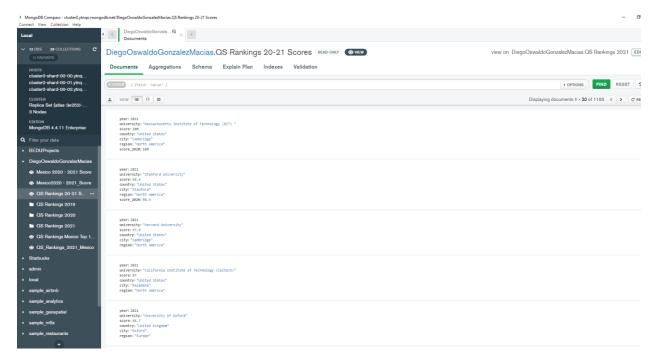
14) ¿Cuál ha sido la diferencia en la posición de México respecto 2020 y 2021?



15) ¿Cuál ha sido la diferencia en calificación o puntaje entre las universidades mexicanas en 2020 y 2021?



4. Vistas relevantes



5. Conclusiones

- Mongo probo ser más fácil de comprender para mí que SQL y eso, junto con el hecho que tiene un desempeño muy bueno, fueron las razones por las que decidí enfocarme en Mongo.
- Las partes mas frustrantes del proyecto fueron: formular las preguntas y generar los JSONs correctos.
- El hecho de que muchas universidades no tengan calificación del año 2020 refleja el impacto que tuvo la pandemia en su decisión de no participar o que QS haya decidido no cambiar las calificaciones.
- El desempeño de las Universidades mexicanas esta muy por detrás, incluso de otras universidades latinoamericanas.
- México debería poner mas interes por mejorar en algunas de las métricas evaluadas para ser mas competitivo.

6. JSONs

MongoDB

```
Pregunta 1
{
filter: {
 rank display: {
 $ne: NaN
},
project: {
 rank_display: 1,
university: 1,
 _id: 0
},
sort: {
rank display: 1
},
limit: 5
        Pregunta 2
filter: {
 rank_display: {
 $ne: NaN
}
},
```

```
project: {
 rank_display: 1,
 university: 1,
_id: 0
},
sort: {
rank_display: -1
},
limit: 5
}
    • Pregunta 3
filter: {
 $and: [
  rank_display: {
  $lte: 30
  }
 },
 region: 'North America'
 ]
},
project: {
 university: 1,
 rank_display: 1,
 region: 1,
 _id: 0
},
limit: 30
}
        Pregunta 4
      filter: {
       $and: [
       {
        rank_display: {
         $Ite: 30
        }
       },
        region: 'Europe'
       ]
      },
      project: {
```

```
university: 1,
   rank_display: 1,
   region: 1,
   _id: 0
  },
  limit: 30
 }
    Pregunta 5
 {
  filter: {
  $and: [
    rank_display: {
    $lte: 30
    }
   },
    region: 'Asia'
  ]
  },
  project: {
   university: 1,
   rank_display: 1,
   region: 1,
  _id: 0
  },
  limit: 30
 }
• Pregunta 6
 [{$match: {
  rank_display: {
  $lte: 10
 }}, {$group: {
  _id: null,
  universities: {
  $sum: 1
  },
  scores: {
  $sum: '$score'
 }}, {$addFields: {
  calificacion_promedio_top10: {
   $divide: [
   '$scores',
   '$universities'
```

```
}
}}]
Pregunta 7
[{$match: {
region: 'Latin America'
}}, {$limit: 10}, {$group: {
_id: null,
universities: {
 $sum: 1
},
scores: {
 $sum: '$score'
}}, {$addFields: {
calificacion_promedio_top10: {
 $divide: [
  '$scores',
  '$universities'
 ]
}
}}]
  Pregunta 8
[{$match: {
region: 'North America'
}}, {$limit: 10}, {$group: {
_id: null,
universities: {
 $sum: 1
},
scores: {
 $sum: '$score'
}}, {$addFields: {
calificacion_promedio_top10: {
 $divide: [
  '$scores',
  '$universities'
 ]
}
}}]
  Pregunta 9
[{$match: {
country: 'Mexico'
}}, {$limit: 10}, {$group: {
_id: null,
```

```
universities: {
 $sum: 1
},
scores: {
 $sum: '$score'
}}, {$addFields: {
calificacion_promedio_top10: {
 $divide: [
 '$scores',
 '$universities'
 ]
}
}}]
  Pregunta 10
[{$match: {
score: {
 $Ite: 47.8
}}, {$group: {
_id: null,
Universidades_superiores_a_Mexico: {
 $sum: 1
}
}}]
 Pregunta 11
[{$match: {
rank_display: {
 $lte: 50
},
country: 'United States'
}}, {$group: {
id: null,
'Frecuencia EUA': {
 $sum: 1
}
}}]
  Pregunta 12
[{$lookup: {
from: 'QS Rankings 2020',
localField: 'rank_display',
foreignField: 'rank_display',
as: 'rank old'
}}, {$addFields: {
rank_20: {
```

```
$arrayElemAt: [
 '$rank_old',
 0
 ]
}
}}, {$addFields: {
rank_display_2020: '$rank_20.rank_display'
}}, {$project: {
_id: 0,
year: 1,
rank_display_2020: 1,
rank_display: 1,
university: 1,
country: 1,
city: 1,
region: 1
}}]
  Pregunta 13
[{$lookup: {
from: 'QS Rankings 2020',
localField: 'score',
foreignField: 'score',
as: 'score_old'
}}, {$addFields: {
score_20: {
 $arrayElemAt: [
 '$score_old',
 0
 ]
}
}}, {$addFields: {
score_2020: '$score_20.score'
}}, {$project: {
_id: 0,
year: 1,
score_2020: 1,
score: 1,
university: 1,
country: 1,
city: 1,
region: 1
}}]
 Pregunta 14
[{$match: {
country: 'Mexico'
}}, {$limit: 10}, {$lookup: {
from: 'QS Rankings 2020',
localField: 'rank_display',
```

```
foreignField: 'rank_display',
as: 'rank_old'
}}, {$addFields: {
rank_20: {
 $arrayElemAt: [
 '$rank_old',
 0
 ]
}
}}, {$addFields: {
rank_2020: '$rank_20.rank_display'
}}, {$project: {
_id: 0,
year: 1,
rank_2020: 1,
rank_display: 1,
university: 1,
country: 1,
city: 1,
region: 1
}}]Pregunta 15
[{$match: {
country: 'Mexico'
}}, {$limit: 10}, {$lookup: {
from: 'QS Rankings 2020',
localField: 'score',
foreignField: 'score',
as: 'score_old'
}}, {$addFields: {
score_20: {
 $arrayElemAt: [
 '$score_old',
 0
 ]
}
}}, {$addFields: {
score_2020: '$score_20.score'
}}, {$project: {
_id: 0,
year: 1,
score_2020: 1,
score: 1,
university: 1,
country: 1,
city: 1,
region: 1
}}]
```