



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA

**Apuntes de SQL para la
asignatura de Inteligencia de
Negocios**

MATERIAL DIDÁCTICO

Que para obtener el título de

Ingeniero Industrial

P R E S E N T A

Rodrigo Reynoso Fuentes

ASESORA DE MATERIAL DIDÁCTICO

Dra. Susana Casy Téllez Ballesteros



Ciudad Universitaria, Cd. Mx., 2021

Agradecimientos

Este trabajo no habría podido ser posible sin la supervisión de mi asesora, por esto, quiero agradecer profundamente a la Dra. Susana C. Téllez Ballesteros por su tiempo, paciencia y guía. Siempre lo valoraré.

Agradezco especialmente al M. en I. Ricardo Torres Mendoza por haberme ayudado en la tarea de adquirir nuevo conocimiento y encaminarme al mundo profesional.

Gracias a la Universidad Nacional Autónoma de México por haberme brindado la oportunidad de aprender de los mejores, así como a los sinodales de este trabajo de titulación: a la M. en I. Silvina Hernández García, la M. en I. Hilda Reyes Solís Vivanco, el Fis. Juan Velázquez Torres, la M. en I. María de Lourdes Arellano Bolio y mi asesora la Dra. Susana C. Téllez Ballesteros.

De igual manera agradezco al Programa de Apoyo a Proyectos para la Innovación y Mejoramiento de la Enseñanza (PAPIME)- Proyecto PE102218 “Elaboración de Material didáctico (notas, presentaciones, series de ejercicio y estudios de caso) y cursos en línea respecto a Simulación de Sistemas Complejos e Inteligencia de Negocios (Business Intelligence)” y a todos los involucrados, esto por permitirme la elaboración de este trabajo que funge como mi proyecto de titulación.

Quiero agradecer a mis dos grandes amistades, Eduardo y Gabriel, por apoyarme en cada decisión que he tomado y estar a un lado en cada paso que doy.

Finalmente agradecer a mi familia por todo el apoyo que me han brindado a lo largo de mi vida, gracias por estar ahí. Gracias a mi madre, por ser la mejor madre que alguien podría tener y siempre procurar de mí, a pesar de todo. Gracias a mi hermano, por ser mi maestro en la vida siempre. Gracias a mi hermana, por enseñarme que la vida sí da más oportunidades y hay que aprovecharlas. Gracias a mi novio, por amarnos y apoyarnos en cada decisión que tomamos. Siempre estaré ahí para ustedes.

Índice

1. Marco Contextual	6
1.1. Teorías del aprendizaje	6
1.1.1. Conductismo	6
1.1.2. Constructivismo	6
1.1.3. Cognitivismo	6
1.2. La enseñanza como proceso	7
1.2.1. Planeación	7
1.2.2. Revisión	7
1.2.3. Evaluación	7
1.2.4. Implementación	8
1.3. Componentes del proceso de aprendizaje	8
1.3.1. Atención	8
1.3.2. Memoria	9
1.3.3. Lenguaje	9
1.3.4. Organización	10
1.3.5. Escritura	11
1.4. Emociones	11
1.5. Enseñanza de la ingeniería	12
1.5.1. Profesorado en la industria	12
1.5.2. Currículum	12
1.5.3. Formalizar el aprendizaje de vida	13
1.5.4. Diversidad	13
1.6. El rol del docente	13
1.6.1. Motivación	14
1.7. El rol del estudiante	14
1.8. Perfil del estudiante	14
1.9. Perfil del docente	15
1.10. Perfil del egresado	16
2. SQL	17
2.1. ¿Qué es SQL?	17
3. Instalar sistema para la gestión de bases de datos	17
3.1. Instalar PGAdmin 4	17
3.2. Instalar PostgreSQL	24

4. Estudio de caso	33
4.1. Especificaciones de la base de datos IMDb	33
4.2. Iniciar PostgreSQL	35
4.3. Iniciar PostgreSQL	36
5. Instrucciones de extracción de SQL	39
5.1. Instrucción SELECT (Seleccionar)	39
5.1.1. Práctica: Instrucción SELECT	40
5.1.2. Ejercicio: Instrucción SELECT	43
5.2. Instrucción SELECT DISTINCT (Selección Distinta)	43
5.2.1. Práctica: Instrucción SELECT DISTINCT	43
5.3. Instrucción SELECT WHERE (Selección Dónde)	44
5.3.1. Práctica: Instrucción SELECT WHERE	45
5.3.2. Ejercicio: Instrucción SELECT WHERE	46
6. Instrucciones de limpieza de SQL	46
6.1. Instrucción COUNT (Contar)	46
6.1.1. Práctica: Instrucción COUNT	47
6.1.2. Ejercicio: Instrucción COUNT	49
6.2. Instrucción ORDER BY (Order Por)	49
6.2.1. Práctica: Instrucción ORDER BY	50
6.3. Instrucción BETWEEN (Entre)	51
6.3.1. Práctica: Instrucción BETWEEN	51
6.4. Instrucción IN (En)	52
6.4.1. Práctica: Instrucción IN	53
6.5. Instrucción LIKE (Como)	53
6.5.1. Práctica: Instrucción LIKE	54
7. Funciones de agregación	57
7.0.1. Práctica: Funciones de agregación	57
7.1. Instrucción GROUP BY (Agrupar por)	61
7.1.1. Práctica: Instrucción GROUP BY	61
7.2. Instrucción HAVING (Teniendo)	62
7.2.1. Diferencia entre HAVING y WHERE	62
7.2.2. Práctica: Instrucción HAVING	62
8. Instrucción AS (Como)	63

9. JOINS (Uniones)	65
9.0.1. Práctica: INNER JOIN	67
9.1. Tipos de JOIN	68
9.1.1. INNER JOIN	68
9.1.2. FULL OUTER JOIN	68
9.1.3. LEFT OUTER JOIN	68
9.1.4. LEFT OUTER JOIN con WHERE	69
9.1.5. FULL OUTER JOIN con WHERE	69
10.Sub-consultas	69
11.Sub-tablas	70
12.Ejercicios de repaso	72
12.1. Instrucción SELECT	72
12.2. Instrucción SELECT DISTINCT	72
12.3. Instrucción SELECT WHERE	72
12.4. Instrucción COUNT	73
12.5. Instrucción ORDER BY	73
12.6. Instrucción BETWEEN	73
12.7. Instrucción IN	74
12.8. Instrucción LIKE	74
12.9. Instrucción GROUP BY	74
12.10Instrucción HAVING	75
12.11Instrucción INNER JOIN	75
12.12Sub-consulta	76
13.Conclusiones	77

Objetivo de los apuntes

Proporcionar al alumno las herramientas necesarias del lenguaje de consulta estructurado (SQL) para poder poner en práctica los conocimientos teóricos de la integración y estructura de una base de datos. Poder conocer comandos básicos de SQL para poder extraer información relevante para el estudio profundo de estos y poder tomar decisiones de negocios estratégicas.

Alcance

El alumno trabajará con una base de datos construida, a partir de esta base de datos podrá hacer uniones de tablas, extracciones y creación de tablas usando los comandos básicos de SQL. Todo con el propósito de generar .csv o .tsv final.

1. Marco Contextual

1.1. Teorías del aprendizaje

Aunque hay diferentes enfoques, hay tres acercamientos básicos de la teoría del aprendizaje: conductismo, constructivismo y cognitivismo. En esta sección se da una breve introducción de cada tipo de teoría de aprendizaje [4].

1.1.1. Conductismo

El conductismo asume que el estudiante es esencialmente pasivo, el cual responde esencialmente a estímulos del ambiente. Se presume que el estudiante comienza como una hoja en blanco y el comportamiento se va moldeando a través de reforzamientos positivos o negativos. Ambos reforzamientos, positivos y negativos incrementan la probabilidad de que el comportamiento anterior se repita de nuevo. Por lo contrario, el castigo decrementa la probabilidad de que el comportamiento anterior suceda de nuevo. Positivo indica la aplicación de un estímulo y negativo indica la retención de un estímulo. Esto define aprender como el cambio del comportamiento en el estudiante [3].

1.1.2. Constructivismo

El constructivismo dice que aprender es un proceso activo y contextualizado de construir conocimiento en vez de adquirirlo. El conocimiento es construido basado en experiencias profesionales e hipótesis del ambiente. Cada persona tiene una diferente interpretación y construcción del proceso del conocimiento. El estudiante no es una hoja en blanco, pero trae experiencias pasadas y factores culturales dentro de una situación [1, 9].

1.1.3. Cognitivismo

El Cognitivismo se enfoca en las actividades mentales internas, abrir la caja negra de la mente humana es necesario para entender como la gente aprende. El conocimiento puede ser visto como un esquema o construcciones mentales simbólicas. El aprendizaje se define como el cambio en los esquemas de un alumno.

A diferencia del conductismo, la gente no responde a estímulos, las personas son seres racionales que requieren de una participación activa para aprender, cuyas acciones son consecuencias el pensamiento [2].

1.2. La enseñanza como proceso

La enseñanza es fundamentalmente un proceso, donde se incluye la planeación, implementación, evaluación y revisión. Planeas y enseñar una lección son ideas familiares de la mayoría de los docentes. Los pasos que se dan más por alto son la evaluación y la revisión. Sin evaluaciones en el aula o algún otro medio de recibir retroalimentación de manera regular, es sorprendentemente fácil malinterpretar si un método o estrategia de enseñanza en particular ha sido eficaz. Un docente puede crear ambientes de confianza al someterse a la retroalimentación de los estudiantes. Los pasos del proceso de enseñanza los dividiremos en planeación, revisión, evaluación e implementación [5].

1.2.1. Planeación

Hay diferentes niveles de establecer objetivos para la enseñanza, desde la escala de un semestre entero hasta una sola clase. Se tiene la tarea fundamental de ayudar a los estudiantes a aprender cómo pensar críticamente y entender los conceptos básicos y herramientas de la disciplina impartida. El docente debe ser capaz de tener objetivos diarios. Como docente, probablemente no se deba diseñar un curso completo, pero se debe pensar cómo los objetivos diarios suman a los objetivos principales del curso [5, 7].

1.2.2. Revisión

Revisar la pedagogía ayudará a los estudiantes a aprender y mantenerse interesados. Si el docente se mantiene enfocado en el aprendizaje de los estudiantes, encontrará un mayor significado al proceso típico de una clase, discusión y examen. Se deben evitar los sentimientos de adversarios, el proceso de aprendizaje incita a una relación de cooperación y descubrimiento mutuo entre docente y alumno. Ernest Boyer ayudó a redefinir la noción de las relaciones escolares al incluir la beca de la enseñanza como actividad culminante del proceso de investigación de descubrimiento, integración y aplicación del conocimiento (Boyer 1990) [5, 7].

1.2.3. Evaluación

Evaluaciones recurrentes de los estudiantes y del docente son esenciales para el éxito del enseñante. Para realmente entender si se está enseñando de manera efectiva y que los estudiantes están aprendiendo de manera efectiva, es crucial que el docente evalúe activa y regularmente lo que sus estudiantes han aprendido. Si el docente es capaz de pedir retroalimentación de los estudiantes regularmente (no sólo al final del semestre), puede modificar y mejorar sus estrategias de enseñanza. Las evaluaciones no tienen por qué ser demasiado complejas. De hecho, cuanto más centrado esté en la evaluación, mayor será el impacto que tendrán sus cambios [5, 7].

1.2.4. Implementación

Los mejores planes no tienen sentido si no son implementados. Sin embargo, la mayor parte del trabajo hecho en la enseñanza se hace durante la planeación y la preparación, muchas grandes ideas jamás son implementadas porque es más fácil seguir haciendo lo mismo. El docente no debe tener miedo en implementar nuevas ideas. Si algo no ha funcionado ¿por qué no cambiarlo y hacer algo nuevo? Al menos que se tenga la iniciativa de cambiar y experimentar, será difícil que el docente mejore sus habilidades de enseñanza [5, 7].

1.3. Componentes del proceso de aprendizaje

La habilidad del docente de llevar a cabo instrucciones diferenciadas como resultado de la evaluación de sus estudiantes al ajustar las estrategias y tácticas, requiere un sofisticado conocimiento y habilidades.

Para poder adquirir dicha habilidad, el docente debe tener un amplio conocimiento de cada uno de los componentes cognitivos del proceso de aprendizaje, cómo lucen cuando están funcionando y cómo los subcomponentes específicos de cada uno lucen. De igual manera, el docente debe desarrollar un repertorio de estrategias de las cuales sabe cuál usar para una tarea en especial, en un momento específico. Usar una gran estrategia en el momento inadecuado no traerá ningún beneficio, sólo frustrará a los estudiantes y docentes.

Existen cinco componentes interactivos del proceso de aprendizaje: atención, memoria, lenguaje, procesamiento, organización y escritura. Estos procesos interactúan no sólo entre sí, sino con emociones, ambiente de clase, comportamiento, habilidades sociales, docentes y familia.

Para poder motivar y enseñar a todos los estudiantes a un nivel óptimo, los docentes deben comprender el proceso de aprendizaje, entender y responder a los perfiles emocionales y cognitivos de cada uno de los estudiantes y seleccionar estrategias de instrucción y tácticas que sean efectivas para diferentes estudiantes [5].

1.3.1. Atención

Poner atención es el primer paso para aprender. Es fácil, para la mayoría de nosotros, poner atención a cosas que son interesantes o emocionantes para nosotros. Es difícil, para la mayoría de nosotros, poner atención cuando las cosas no nos resultan interesantes o emocionantes. Cuando algo no nos parece interesante, es más fácil distraerse, pasar a un tema o

actividad más estimulante, o desconectarse.

La tarea del docente es construir clases que conecten con los estudiantes. Relacionar lo que es enseñado con la vida cotidiana de los estudiantes puede ayudar.

El movimiento físico puede ayudar a despertar la mente. Cuando los estudiantes muestran señales de falta de atención o inquietud, los maestros pueden proporcionar al estudiante oportunidades para moverse. Varios estudiantes con problemas de atención necesitan moverse para mantenerse activos.

1.3.2. Memoria

La memoria es un proceso complejo que usa tres sistemas para ayudar a una persona a recibir, usar, almacenar y recuperar información. Los tres sistemas de la memoria son:

1. Memoria a corto plazo.
2. Memoria en proceso.
3. Memoria a largo plazo.

Es importa recordar que, el que el estudiante entienda algo, no necesariamente garantiza que lo recordará. Con el fin de aumentar la probabilidad de que todos los estudiantes elaboren nueva información, los docentes deben activar su conocimiento previo y hacer que la nueva información sea significativa para ellos.

Los estudiantes que tiene dificultades en la memoria a corto plazo y memoria en proceso necesitarán que las instrucciones les sean repetidas. Dar indicaciones de manera oral y escrita, y dar ejemplos de lo que es esperado, ayudará a los estudiantes. A estos se les debe pedir identificar la información importante, formular preguntas tipo examen y responderlas. Esta táctica también es efectiva en grupos de aprendizaje cooperativos [5].

1.3.3. Lenguaje

El lenguaje es el principal medio por el cual damos y recibimos información en la universidad. Los dos sistemas de procesamiento del lenguaje son expresivo y receptivo. Usamos el lenguaje expresivo cuando hablamos y escribimos, y usamos el lenguaje receptivo cuando leemos y escuchamos. Los estudiantes con buenas habilidades del procesamiento del lenguaje usualmente les va bien en la universidad. Los problemas con el lenguaje, por otro lado, pueden

afectar la habilidad del estudiante de comunicarse efectivamente, entender y almacenar información verbal y escrita, entender lo que los demás dicen, y mantener relaciones con los demás.

La mayoría de los estudiantes, especialmente aquellos con debilidades en el lenguaje escrito, se beneficiarán de usar un procedimiento de escenificación para la escritura expositiva y creativa.

Todos los estudiantes se benefician de una enseñanza sistemática, cumulativa y explícita basada en la lectura y la escritura. Los estudiantes que tienen desafíos de lenguaje receptivo como una velocidad de procesamiento más lenta deben usar mucha energía mental para escuchar y, por lo tanto, pueden cansarse fácilmente. Las instrucciones orales también pueden necesitar ser repetidas o proporcionadas por escrito.

Ampliando la manera en la que comunicamos información en el salón de clases puede conectar a todos los estudiantes más al tópico de clase, especialmente a estudiantes con problemas del lenguaje. Usar comunicaciones visuales como los son imágenes y videos para reforzar la comunicación verbal es de ayuda para todos los estudiantes, en especial para los estudiantes con problemas receptivos del lenguaje [5].

1.3.4. Organización

Procesamos y organizamos información en dos maneras principales: simultánea (espacial) y sucesiva (secuencial). El procesamiento simultáneo es el proceso que usamos para ordenar y organizar la información en el espacio. Tener un buen sentido de orientación y ser capaz de visualizar como embonan las piezas de un rompecabezas son ejemplos del procesamiento simultáneo. Procesamiento sucesivo es lo que se usa para ordenar y organizar la información en tiempo y secuencia. Conceptos del tiempo, fechas y el orden de las oraciones en los párrafos son ejemplos de procesamiento secuencial. Los estudiantes que son buenos en la organización sucesiva, por lo general, tienen pocos o ningún problema con la administración del tiempo y, generalmente, les resulta fácil organizar un ensayo en una secuencia que sea lógica.

Los estudiantes que tienen complicaciones al entender problemas espaciales y geográficos pueden necesitar explicaciones verbales adicionales. Pueden beneficiarse de la redacción de explicaciones y descripciones escritas de la información contenida en cuadros, gráficos o diagramas. Los maestros deben modelar este proceso para todos los estudiantes.

Los estudiantes que tienen problemas en recordar secuencias de información pero que son fuertes en el procesamiento simultáneo, deberían beneficiarse de los organizadores gráficos, y

de hacer diagramas o diagramas de flujo de información secuencial como eventos en la historia en lugar de la línea de tiempo estándar. Pueden beneficiarse de programas de software que organizan conceptos e información en mapas visuales.

Practicar el aprendizaje cooperativo permite que las fortalezas de procesamientos y organización de cada estudiante sea usado en beneficio del grupo. Por ejemplo, aquellos que son fuertes en la organización simultánea pueden crear el gráfico, el visual o el mapa del grupo, y aquellos que son fuertes en la organización sucesiva pueden ser los organizadores de los pasos de la tarea, los capataces, los cronometradores y los que marcan el ritmo [5].

1.3.5. Escritura

El proceso de la escritura requiere coordinación neuronal, visual y muscular para producir trabajos escritos. No es un acto de voluntad sino un acto de coordinación entre todas las funciones. Regularmente el estudiante que parece desmotivado para completar trabajo escrito es el estudiante cuya coordinación de escritura es torpe. De manera similar, como la práctica no hará a un futbolista una estrella a partir de un torpe en el deporte, la práctica y los actos no harán a una estrella a alguien cuyo el cableado neurológico no le permite ser un atleta escritor de alto rendimiento.

Los estudiantes con dificultades de escritura a mano pueden beneficiarse de la oportunidad de proporcionar respuestas orales a ejercicios, pruebas y exámenes [5].

1.4. Emociones

Las emociones controlan el interruptor de encendido y apagado para el aprendizaje. Cuando se está relajado y calmado, el procesamiento está encendido. Cuando se está estresado, ansioso o con miedo, el procesamiento está apagado. En el salón de clases, la tensión mantiene la mente cerrada. El crear un ambiente libre de amenazas o clima donde los errores son bienvenidos como oportunidades de aprendizaje reduce la tensión, abre la mente e incrementa la oportunidad de aprendizaje.

Entre más sepa el docente sobre cómo la información es procesada, manipulada y creada, más se dará cuenta cuando algo está funcionando y cómo luce cuando comienza a descomponerse en pequeñas partes. Entonces, en vez de pensar que un estudiante está desmotivado, los docentes buscarán si la atención, memoria, lenguaje, organización y escritura es lo que necesita rectificación [5].

1.5. Enseñanza de la ingeniería

La ingeniería no es sólo una ciencia o simplemente una ciencia aplicada. Aunque la ciencia es analítica en el sentido que se esfuerza por comprender la naturaleza, o lo que es, la ingeniería es sintética en el sentido que se esfuerza en crear. La creatividad de la ingeniería está limitada por la naturaleza, costos, preocupaciones de seguridad, impacto ambiental, ergonomía, replicabilidad, manufactura, entre otras cosas.

Mientras el mundo se vuelve más complejo, los ingenieros deben apreciar las dimensiones de la tecnología, tener un conocimiento de los temas globales, ser sensibles a la diversidad cultural y saberse comunicar efectivamente.

Estos imperativos influyen fuertemente en cómo un ingeniero debe ser educado, lo que quiere decir que el estudiante requiere un tipo de educación que está actualmente disponible en la Facultad de Ingeniería de la UNAM, se observan 4 áreas importantes a reformar: Profesorado en la industria, currículum, formalizar el aprendizaje de vida y diversidad [6].

1.5.1. Profesorado en la industria

Los docentes de ingeniería son, en gran escala, juzgados por sus méritos científicos dentro de la facultad y la práctica profesional ingenieril no suele ser uno de esos méritos. El facultado ingenieril actual tienden a ser investigadores capaces, pero muy pocos no están familiarizados con los problemas mundiales, simplemente por el hecho de que jamás ha practicado ingeniería.

Cuando la Academia Nacional de Ingeniería de Estados Unidos convocó al gobierno y el mundo académico en enero de 2002 a 28 líderes de la industria para discutir la investigación sobre la enseñanza y el aprendizaje de la ingeniería, los participantes en el retiro estuvieron de acuerdo en que, aunque es muy necesario centrarse más en las actividades académicas de enseñanza y aprendizaje de la ingeniería, el actual sistema de recompensas del profesorado debe valorar más la experiencia profesional [6].

1.5.2. Currículum

La debilidad del profesorado en la práctica de la ingeniería provoca una gran diferencia entre lo que se enseña en la escuela y lo que se espera de los jóvenes ingenieros por parte de sus empleadores y clientes. Mientras que la física y las matemáticas continuas fueron de gran importancia durante la mayor parte del siglo XX, ahora hay fundamentos adicionales. Por ejemplo, las matemáticas discretas (esenciales para la tecnología digital de la información), las ciencias químicas y biológicas, y el conocimiento de los contextos culturales y empresariales

globales para el diseño son ahora partes importantes del repertorio de un ingeniero [6].

1.5.3. Formalizar el aprendizaje de vida

Se ha dicho que la vida media de los conocimientos de ingeniería (el tiempo en que la mitad de lo que un ingeniero sabe se vuelve obsoleto) está en el rango de dos a ocho años. Esto significa que el aprendizaje permanente es esencial para mantenerse al día a lo largo de toda la carrera de ingeniería, que puede durar unos 40 años. Los fundamentos que aprendiste en la universidad siguen siendo fundamentales, pero no son los únicos en esta profesión que cambia rápidamente [6].

1.5.4. Diversidad

Aunque la diversidad en la Facultad de Ingeniería ha mejorado en los últimos años, menos del 20 por ciento de los estudiantes de primer año que ingresan son mujeres. Entre los profesores de ingeniería de la nación, las cifras son peores, menos del 10 por ciento son mujeres. No sólo como una cuestión de equidad, es un problema de la fuerza laboral y, lo que es más importante, es un problema de calidad. El campo creativo se ve privado de un amplio espectro de experiencias de vida que influyen directamente en un buen diseño de ingeniería [6].

1.6. El rol del docente

El docente no sólo debe proporcionar información, sino ser un mediador entre el alumno y el ambiente, ser el guía del aprendizaje como acompañante del alumno. Para ello, Brooks y Brooks (1999), proponen que para satisfacer la diversidad y tomar ventaja que proporciona dicha diversidad dentro del salón de clases, el docente debe cumplir una serie de características. El perfil del docente puede adecuarse a cualquier nivel de enseñanza-aprendizaje.

El docente estimula la autonomía, iniciativa y liderazgo que ayudan al estudiante a establecer conexiones entre ideas y conceptos, que le permiten plantearse problemas y buscar soluciones. Utiliza una gran diversidad de materiales interactivos donde se presenta a los estudiantes posibilidades reales y luego les ayuda a generar abstracciones. Averigua cómo han comprendido sus alumnos los conceptos antes de compartir con ellos su propia comprensión de los mismos, pues si se les da el conocimiento ya hecho, se les está poniendo en sus manos el último eslabón de un proceso de pensamiento que sólo ellos pueden construir.

El docente es flexible en el diseño de la clase, permite que los intereses y las respuestas de los alumnos orienten el rumbo de las sesiones, determinen las estrategias de enseñanza

y alteren el contenido y se enfoca en aprovechar los momentos en que los estudiantes se muestran más receptivos para ciertos temas y así poder profundizar en ellos. Estimula a los alumnos a entrar en diálogo tanto con el maestro como entre ellos y a trabajar colaborativamente. Promueve el aprendizaje por medio de preguntas inteligentes y abiertas y anima a los estudiantes a que se pregunten entre ellos; dichas preguntas complejas e inteligentes retan a los alumnos a indagar más allá de lo aparente, a profundizar, a buscar respuestas novedosas.

Busca que los alumnos elaboren sus respuestas iniciales. Las respuestas iniciales son un motor que estimula a los alumnos a estructurar y reconceptualizar. Provee tiempo para que los estudiantes construyan hipótesis y las comprueben, hagan relaciones y creen metáforas. El maestro debe crear el ambiente de aprendizaje y permitir a los estudiantes construir y descubrir [6].

1.6.1. Motivación

Es el trabajo de todo docente motivar a cada estudiante. Aprender más sobre el cerebro y el desarrollo de la mente, estudiar nueva información del aprendizaje, hacer el aprendizaje significativo y aprender sobre el aprendizaje, observar el proceso del aprendizaje, monitorear de manera cercano las rupturas y celebrar los éxitos de cada estudiante.

1.7. El rol del estudiante

La responsabilidad de los estudiantes ocurre cuando los estudiantes toman un papel activo en su aprendizaje y reconocen que son responsables de su éxito académico. La responsabilidad de los estudiantes es demostrada cuándo los estudiantes toman decisiones y acciones que los llevan hacia sus objetivos académicos.

Los estudiantes responsables demuestran en su comportamiento integridad y honestidad académica, atienden y participan en clase, completan las tareas dentro de tiempo y calidad, evitan las excusas por su comportamiento, se comunican de manera cuidadora y respetuosa hacia profesores y el resto de la comunidad, son aprendedores comprometidos que dedican el tiempo suficiente de trabajo fuera de las instalaciones educativas y usan los recursos de la universidad y buscan ayuda cuando la necesitan [6].

1.8. Perfil del estudiante

El estudiante en la licenciatura de Ingeniería Industrial en la Facultad de Ingeniería de la UNAM, puede ser egresado de la Escuela Nacional Preparatoria, del Colegio de Ciencias

y Humanidades o de otros programas de Educación Media Superior.

Posee conocimientos de matemáticas en álgebra, geometría analítica y cálculo diferencial e integral de funciones de una variable; también cuenta con conocimientos de física, particularmente en lo que respecta a temas relacionados con mecánica clásica, así como conocimientos generales de química y de computación. De igual manera posee conocimientos de inglés, por lo menos a nivel de comprensión de textos.

Cuenta con alto sentido de responsabilidad, disciplina, interés por el estudio, criterio de decisión y habilidad para las relaciones humanas. En lo que concierne a datos estadísticos, en los últimos 5 años el promedio se refleja de la siguiente manera (<http://www.estadistica.unam.mx/perfiles/>): El 90 % de los alumnos tiene entre 18 y 22 años. 85 % son hombre y 15 % mujeres. El 100 % de ellos es soltero. El 55 % proviene de un bachillerato público, mientras que el 45 % proviene de bachillerato privado. El 48 % de ellos tiene de 1 a 3 hermanos. El 88 % de ellos tiene de 2 a 5 dependientes económicos. El 42 % de ellos trabaja. El 14 % de ellos tiene un ingreso menos a 2 salarios mínimos y el 43 % tienen un ingreso de 2 a 6 salarios mínimos mensuales.

Los datos anteriores demuestran que la mayoría de ellos son jóvenes solteros, cuyo estudio anterior es el de bachillerato. Gran parte de ellos trabaja y forman parte del sustento económico de familias de tamaño promedio [8].

El docente debe ser consciente que trabaja con alumnos de diversas clases sociales y económicas cuyas necesidades prioritarias cambian drásticamente, el docente deberá encontrar la forma de integrar esa diversidad, incluirla y sacar el mayor provecho de ésta.

1.9. Perfil del docente

Los docentes son egresados de carreras de Ingeniería, Matemáticas, Física o carreras cuyo contenido en el área de matemáticas sea similar preferentemente, cuenta con especialidad en el área, así como con experiencia profesional. Es recomendable que cuente con Maestría en Ingeniería, en Ciencias o con un Doctorado.

Impartir clases no sólo requiere de ser experto en el área, también la formación en docencia es importante. La UNAM brinda apoyo constante a través de cursos, talleres y diplomados donde se motiva, actualiza y enseña a los docentes a implementar técnicas de enseñanza.

1.10. Perfil del egresado

Los ingenieros industriales se centran en cómo hacer el trabajo de la manera más eficiente, equilibrando muchos factores, como el tiempo, el número de trabajadores necesarios, la tecnología disponible, las acciones que los trabajadores deben tomar, el logro del producto final sin errores, la seguridad de los trabajadores, las preocupaciones ambientales y el costo. Para encontrar formas de reducir los residuos y mejorar el rendimiento, los ingenieros industriales primero estudian cuidadosamente los requisitos del producto. Luego utilizan métodos y modelos matemáticos para diseñar sistemas de fabricación e información que satisfagan esos requisitos de la manera más eficiente posible.

Los ingenieros industriales utilizan conocimientos y habilidades especializadas en las ciencias matemáticas, físicas y sociales, junto con los principios y métodos de análisis y diseño de ingeniería, para especificar, predecir y evaluar los resultados obtenidos de los sistemas y procesos. A partir de estos resultados, son capaces de crear nuevos sistemas, procesos o situaciones para la coordinación útil de personas, materiales y máquinas y también mejorar la calidad y productividad de los sistemas, tanto físicos como sociales [10].

(http://www.ingenieria.unam.mx/programas_academicos/licenciatura/industrial.php)

2. SQL

2.1. ¿Qué es SQL?

SQL significa Structured Query Language (Lenguaje de consulta estructurado). SQL se utiliza para comunicarse con una base de datos. Según ANSI (American National Standards Institute), es el lenguaje estándar para los sistemas de gestión de bases de datos relacionales. Las sentencias SQL se utilizan para realizar tareas como actualizar datos en una base de datos o recuperar datos de una base de datos. Algunos sistemas comunes de gestión de bases de datos relacionales que utilizan SQL son: Oracle, Sybase, Microsoft SQL Server, Access, Ingres, etc. Aunque la mayoría de los sistemas de bases de datos utilizan SQL, la mayoría de ellos también tienen sus propias extensiones propietarias que normalmente sólo se utilizan en su sistema. Sin embargo, los comandos SQL estándar como *Select*, *Insert*, *Update*, *Delete*, *Create* y *Drop* pueden utilizarse para realizar casi todo lo que se necesita hacer con una base de datos.

3. Instalar sistema para la gestión de bases de datos

Para poder realizar dichas consultas, en este curso utilizaremos uno de los varios softwares libres disponibles en el mercado. Los pasos de instalación descritos en estos apuntes son indicados para la instalación en Sistema Operativo Windows 10.

3.1. Instalar PGAdmin 4

A continuación, los pasos para descargar PGAdmin 4, éste será la interfaz gráfica que se usará para poder ejecutar PostgrSQL. Asegúrese de tener una conexión a internet estable.

1. Ir a la siguiente dirección: <https://www.pgadmin.org/>
2. Dar clic en “Descarga tu copia aquí”, como se muestra en la Figura 1.

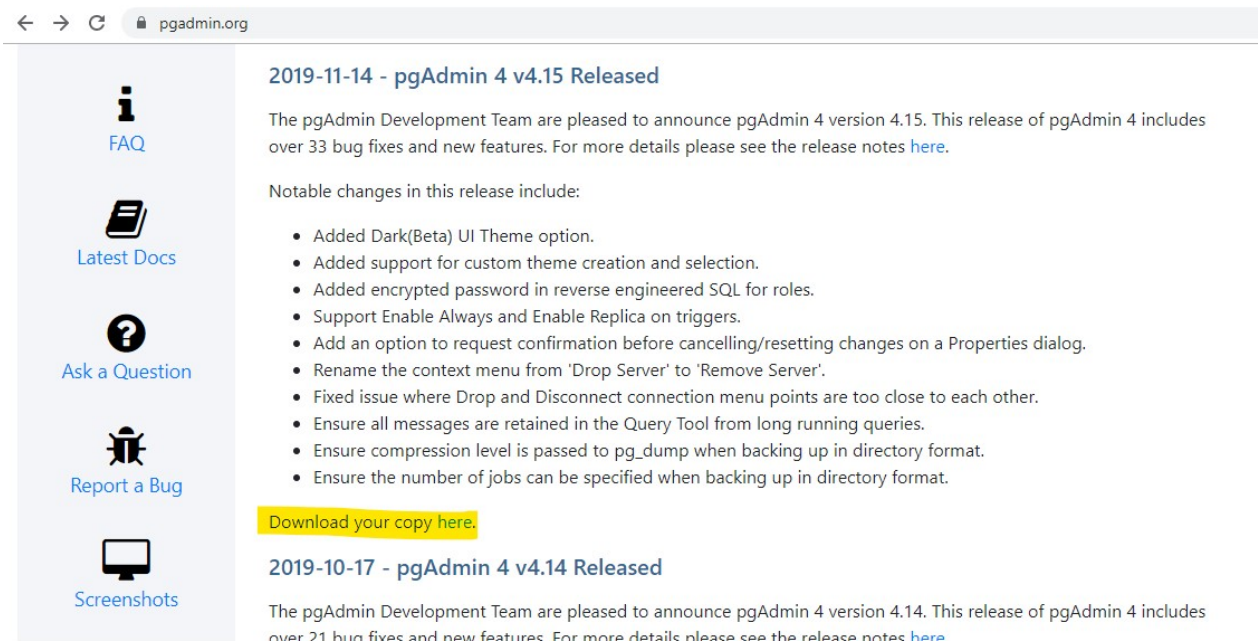


Figura 1: Página de descarga de pgAdmin. Fuente: <https://www.pgadmin.org/>

3. Seleccionar la descarga para el sistema operativo Windows, como se muestra en la Figura 2.

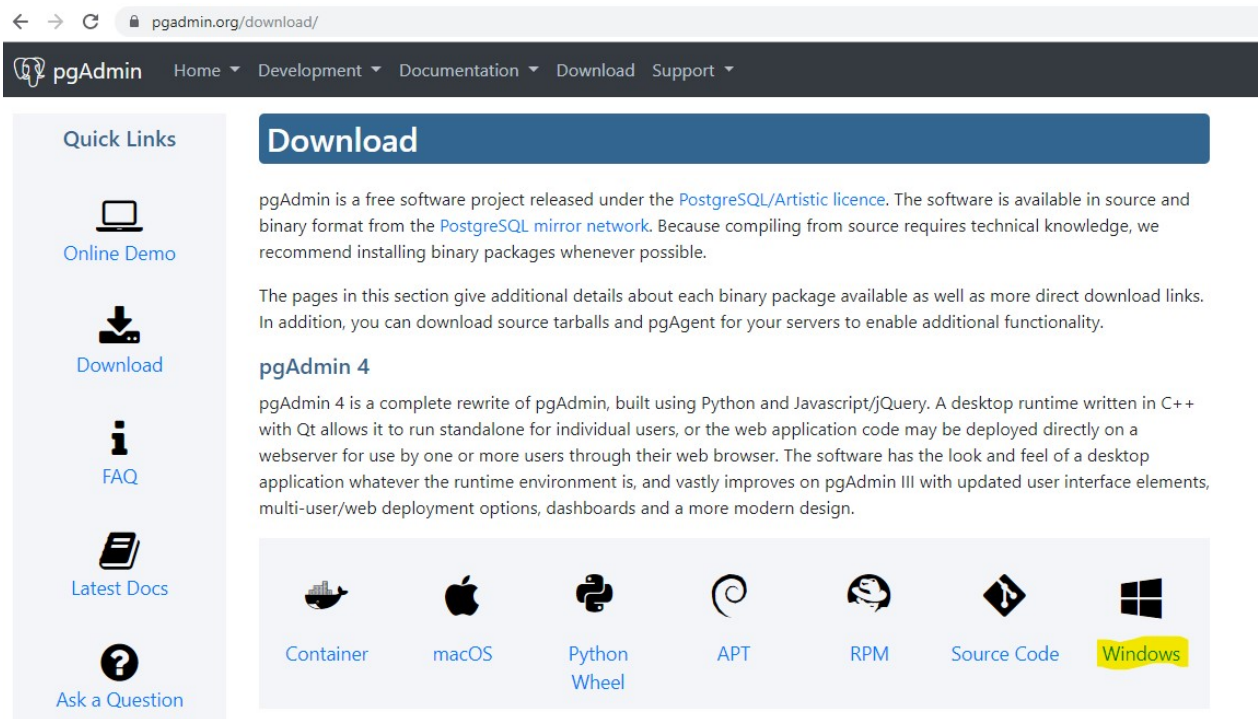


Figura 2: Página dónde se selecciona el sistema operativo de la descarga. Fuente: <https://www.pgadmin.org/>

4. En la ventana que se abre, escoger la última versión disponible pues es la que tendrá el mayor número de actualizaciones y correcciones, como se muestra en la Figura 3.

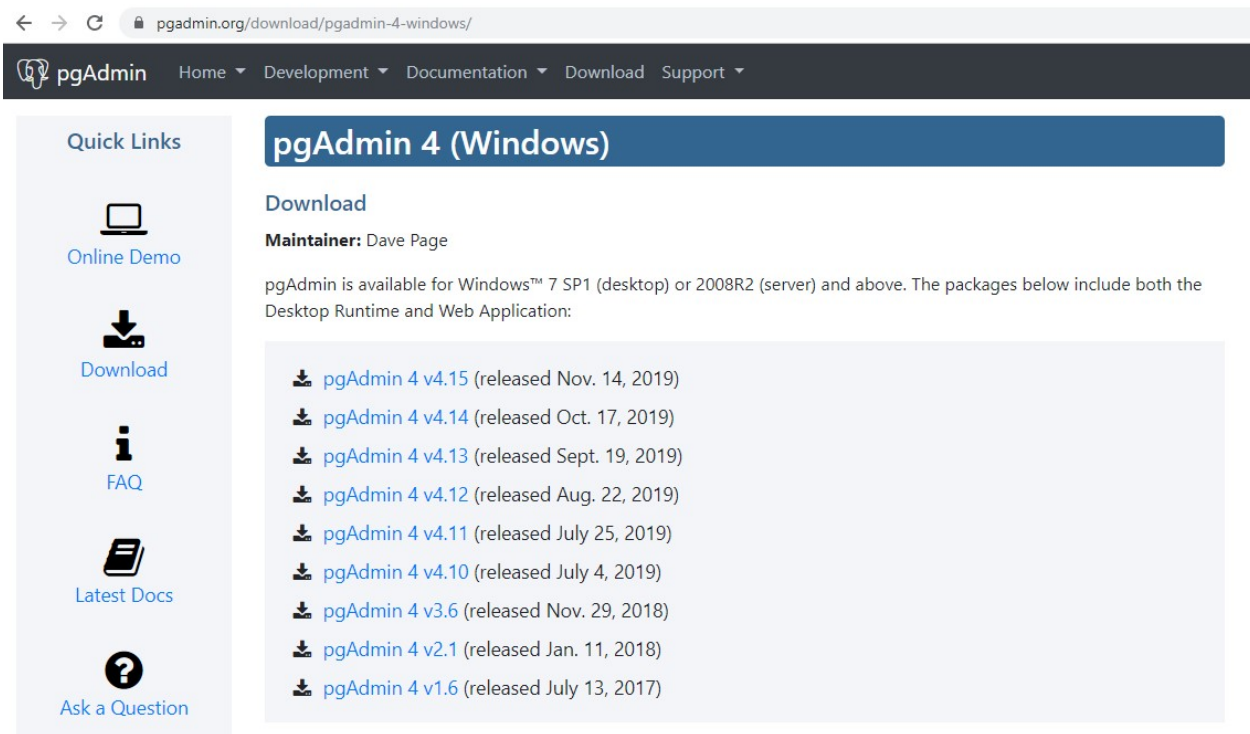


Figura 3: Página dónde se muestran todas las versiones disponibles del software. Fuente:

<https://www.pgadmin.org/>

5. En la página que se abre, seleccionar el archivo ejecutable y comenzará a descargarse, como se muestra en la Figura 4.

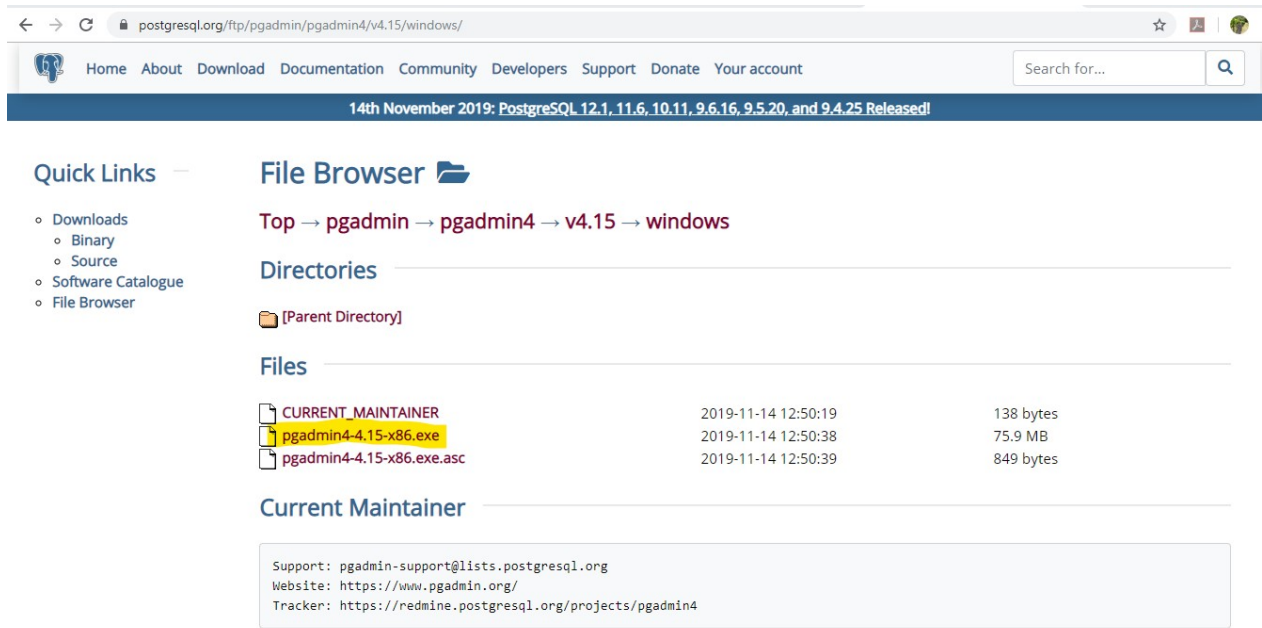


Figura 4: Página dónde se descarga la versión ejecutable. Fuente: <https://www.pgadmin.org/>

6. Verificar la dirección donde se aloja el archivo ejecutable (*.exe).

Una vez descargado y ubicado el archivo ejecutable, siga los siguientes pasos para la instalación:

I Seleccionar “Acepto el acuerdo” y dar clic en “Siguiente”, como se muestra en la Figura 5.

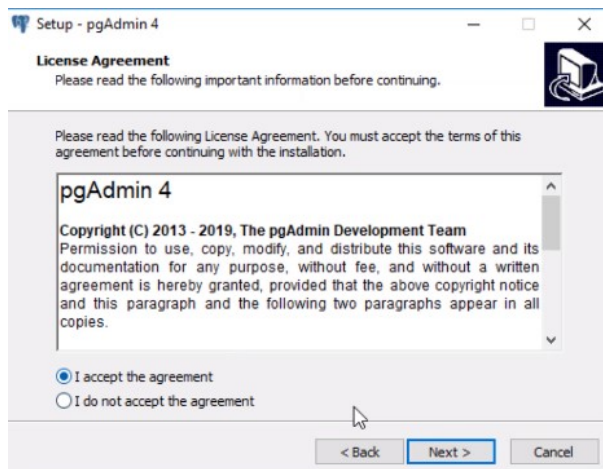


Figura 5: Inicialización del instalador de pgAdmin.

II Dejar la dirección de instalación predeterminada y dar clic en “Siguiente”, como se muestra en la Figura 6.

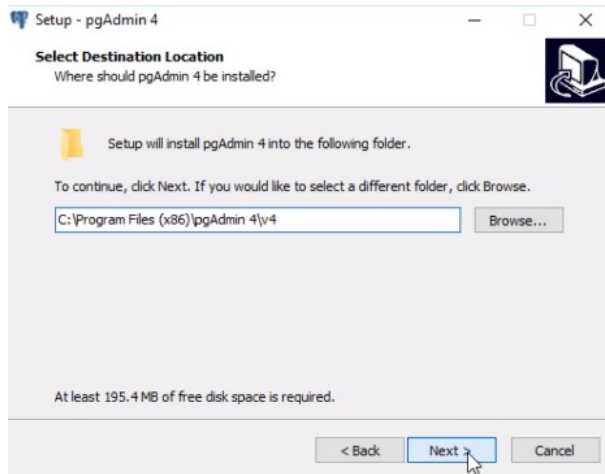


Figura 6: Ventana de verificación de la dirección de instalación.

III Dejar el nombre de la carpeta predeterminado y dar clic en “Siguiente”, como se muestra en la Figura 7.

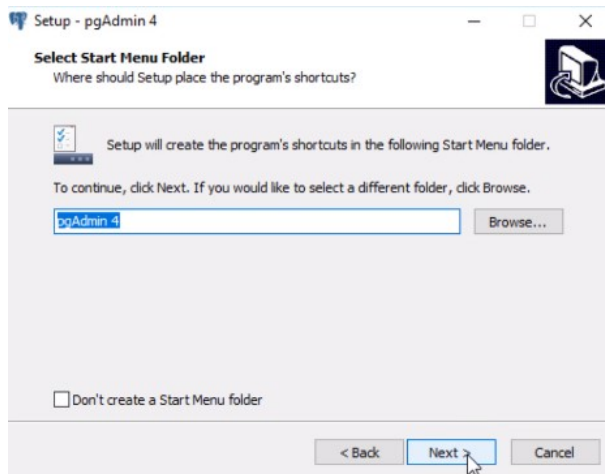


Figura 7: Ventana de verificación del nombre de la carpeta de instalación.

IV Dar clic en “Instalar”, como se muestra en la Figura 8.

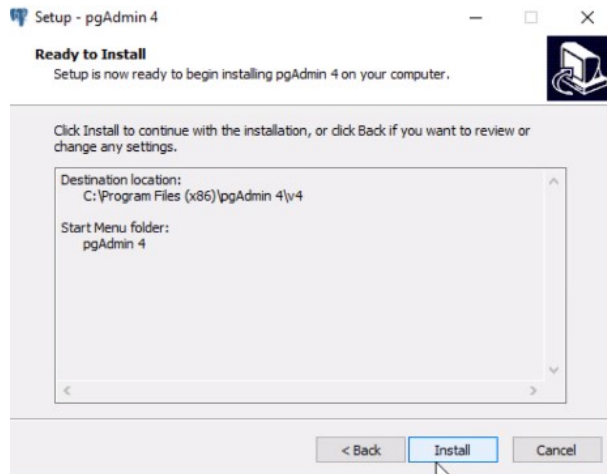


Figura 8: Ventana de confirmación de instalación.

v Esperar a que termine la instalación, como se muestra en la Figura 9.

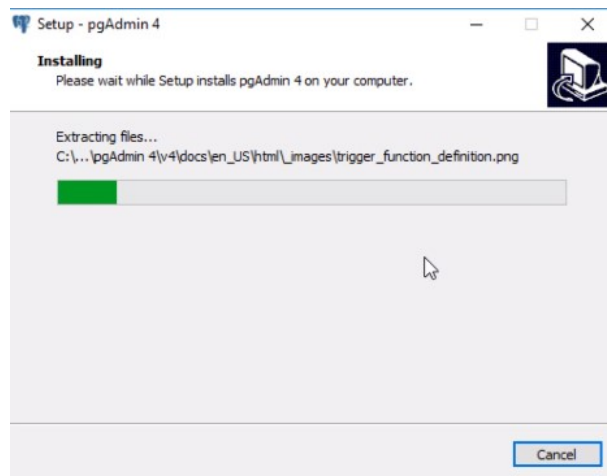


Figura 9: Ventana de avance de la instalación.

vi Cuando haya terminado la instalación y visto que aún se debe instalar otro software, se debe quitar la selección de “Ejecutar pgAdmin ” y dar clic en “Finalizar”, como se muestra en la Figura 10.

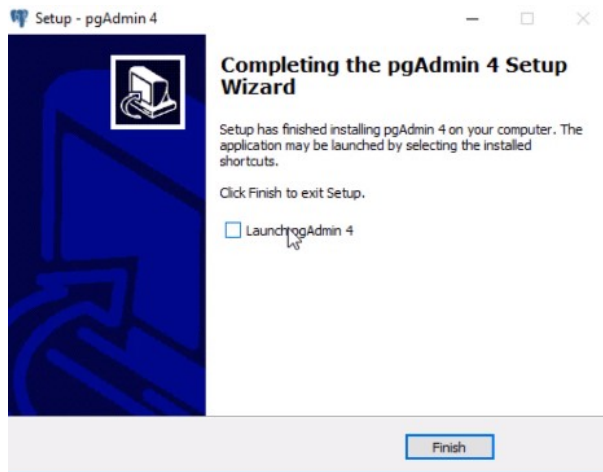


Figura 10: Ventana de finalización de instalación.

3.2. Instalar PostgreSQL

PostgreSQL es la herramienta de consulta que estaremos usando para este curso. Para instalar el software, asegúrese de tener una conexión a internet estable. Siga los siguientes pasos para la instalación:

1. Ir a la dirección: <https://www.postgresql.org/>, como se muestra en la Figura 11.
2. Dar clic en el botón de “Descarga”.

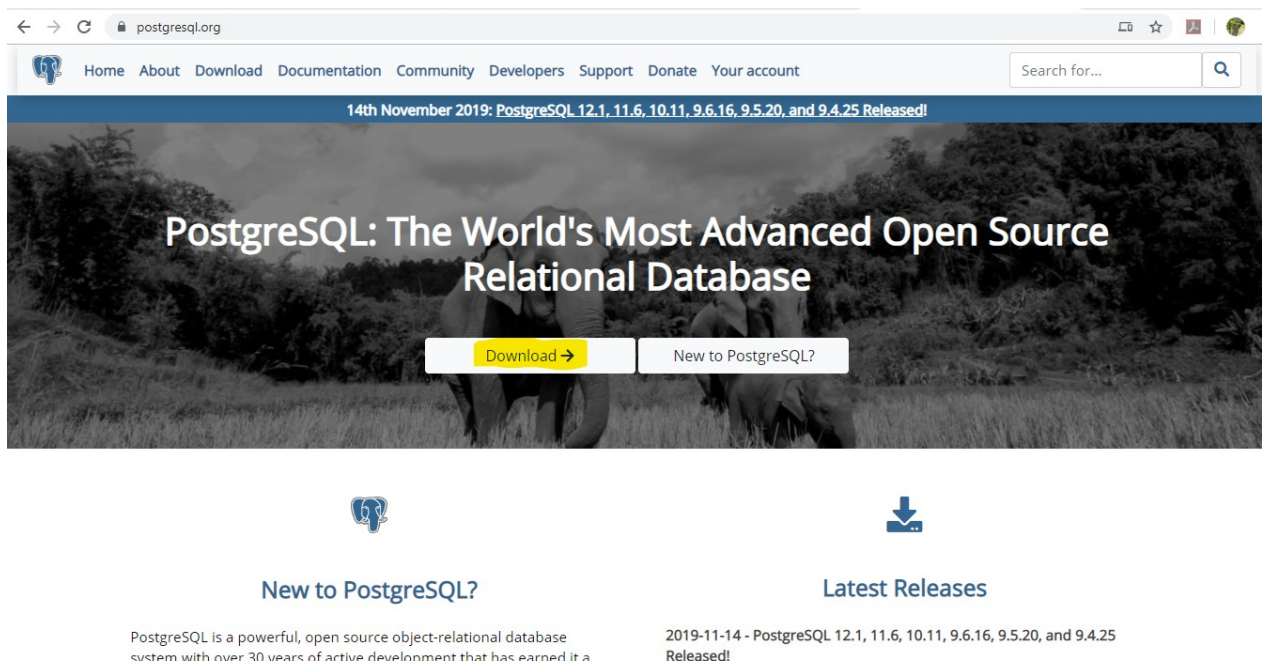


Figura 11: Página de inicio de PostgreSQL. Fuente: <https://www.postgresql.org/>

3. Buscar la subsección de Windows en la sección de “Paquetes Binarios” y dar clic sobre “Windows”, como se muestra en la Figura 12.

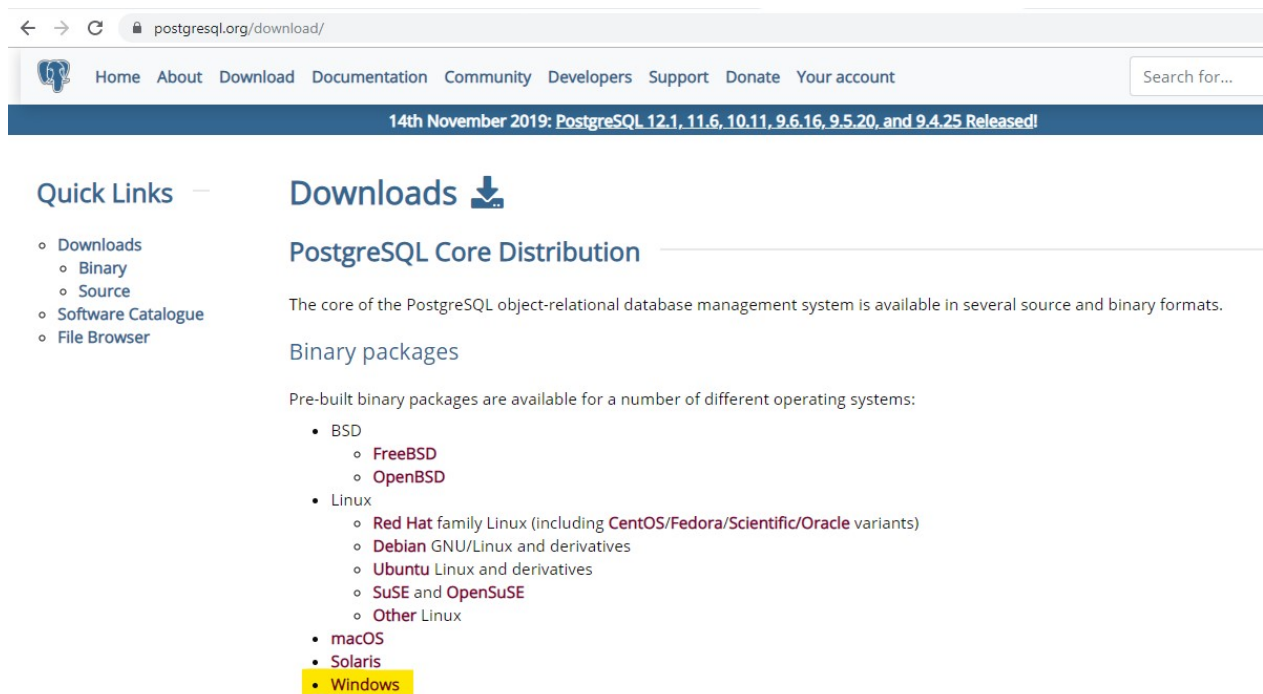


Figura 12: Página de descarga compatible con el sistema operativo del equipo. Fuente:

<https://www.postgresql.org/>

4. En la sección de “Instaladores de Windows”, asegurarse de dar clic en “Descargar el Instalador” del Instalador Interactivo de EnterpriseDB, como se muestra en la Figura 13.

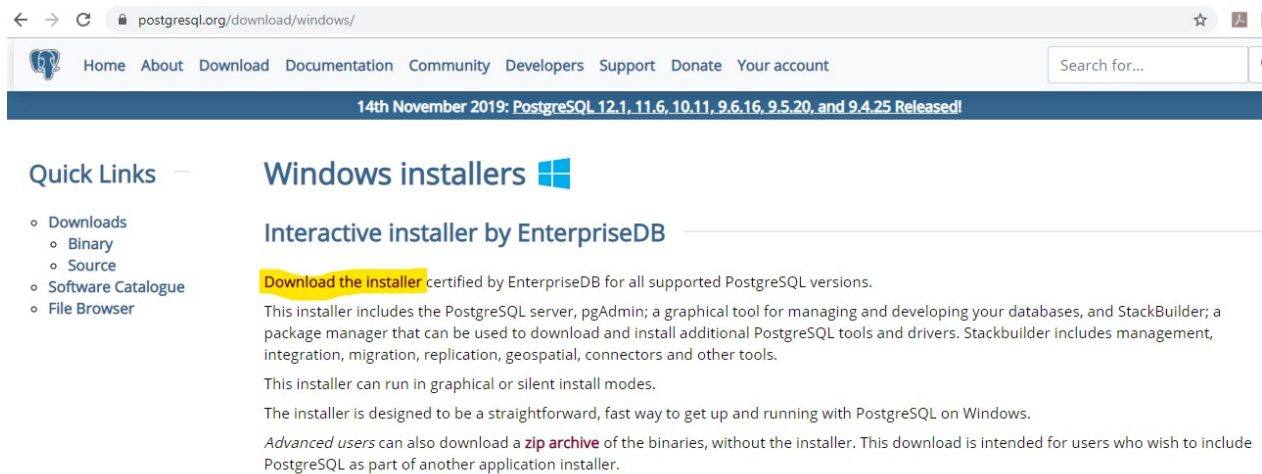


Figura 13: Página de redirección a la descarga del instalador correspondiente. Fuente: <https://www.postgresql.org/>

5. En la siguiente página, descargar la última versión para Windows, como se muestra en la Figura 14.

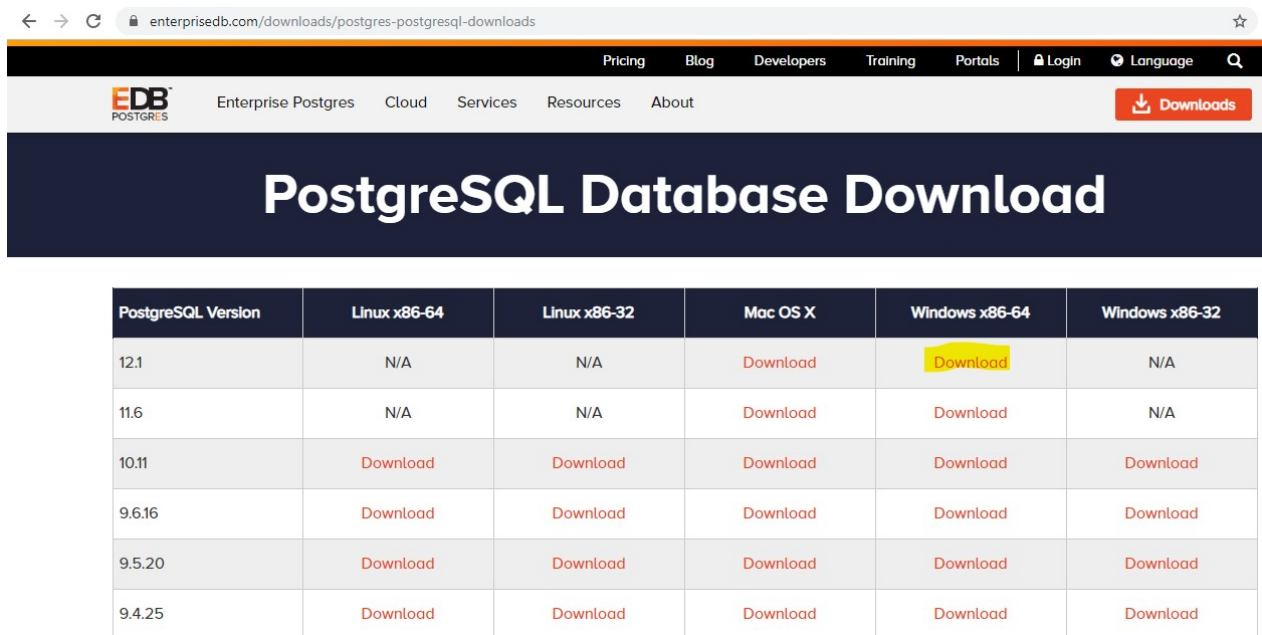


Figura 14: Página de descarga de los instaladores disponibles de Postgre SQL. Fuente: <https://www.postgresql.org/>

Una vez descargado y ubicado el archivo ejecutable, siga los siguientes pasos para la instalación:

I Ejecutar el archivo y dar clic en “Siguiente”, como se muestra en la Figura 15.

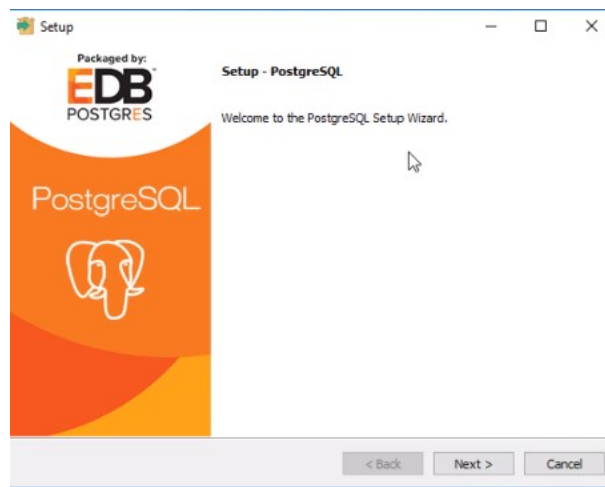


Figura 15: Ventana de inicio de instalación de PostgreSQL.

II Dejar la dirección de instalación predeterminada y dar clic en “Siguiente”, como se muestra en la Figura 16.

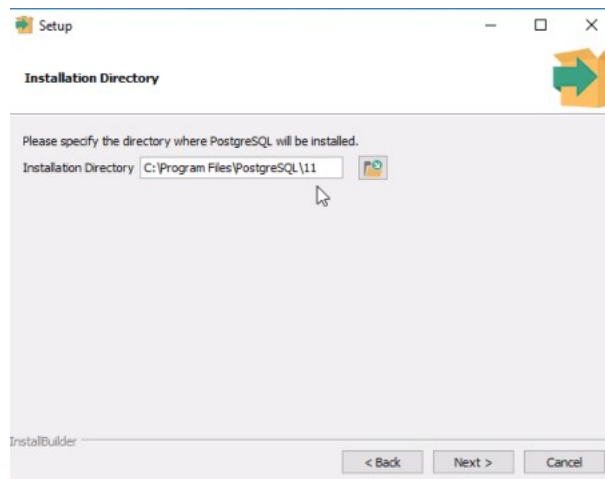


Figura 16: Ventana de verificación de dirección de instalación.

III En la siguiente lista de opciones a instalar, sólo seleccionar PostgreSQL Server, Stack Builder y Command Line Tools. Posteriormente dar clic en “Siguiente”, como se muestra en la Figura 17.

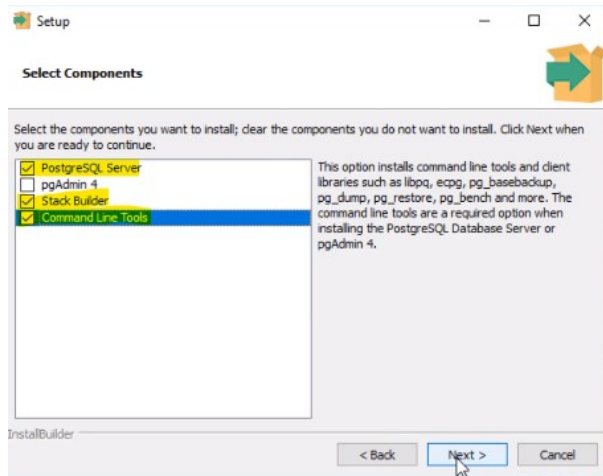


Figura 17: Ventana de selección de objetos a ser instalados.

iv Dejar la dirección predeterminada para PostgreSQL Server y dar clic en “Siguiete”, como se muestra en la Figura 18.

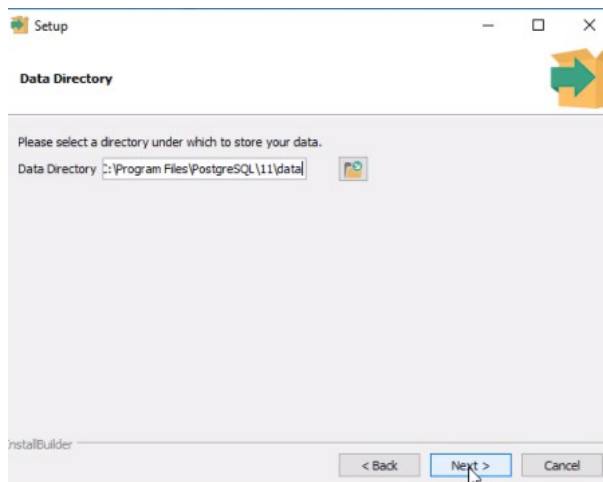


Figura 18: Ventana de verificación de dirección de instalación de PostgreSQL.

v Crear contraseña de acceso para el software y dar clic en “Siguiete”, como se muestra en la Figura 19. **IMPORTANTE:** Crear una contraseña que se pueda recordar fácilmente, en caso de olvidar la contraseña, se deberá borrar PGAdmin y PostgreSQL del equipo y volver a iniciar el proceso de instalación; todo trabajo o avance será perdido en este caso.

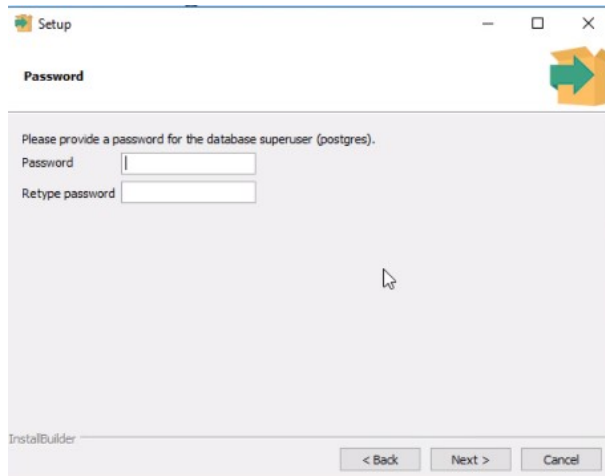


Figura 19: Ventana de creación y verificación de contraseña.

VI Mantener el puerto predeterminado '5432' y dar clic en "Siguiete", como se muestra en la Figura 20.

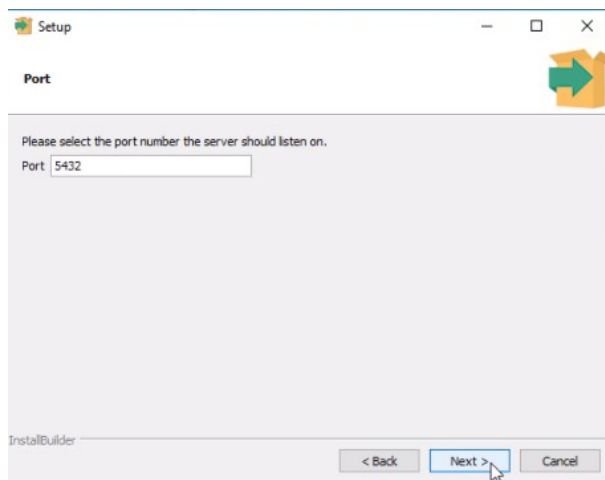


Figura 20: Ventana de verificación de puerto.

VII El local para la instalación debe ser "Local Predeterminad" y dar clic en "Siguiete", como se muestra en la Figura 21.

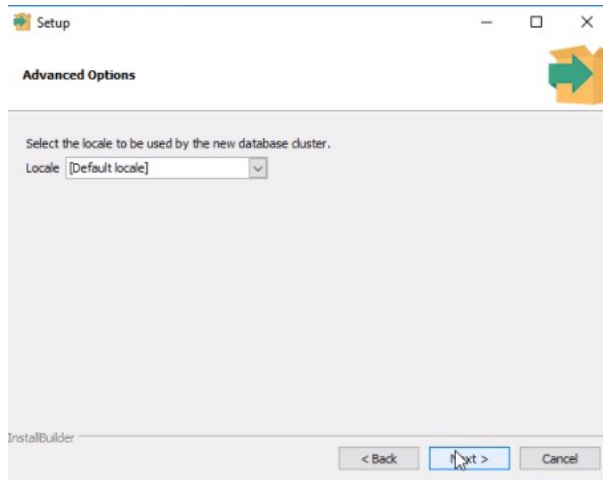


Figura 21: Ventana de selección de local para ser usado para el nuevo grupo de la base de datos.

VIII Una vez verificadas las configuraciones de instalación, dar clic en “Siguiete”, como se muestra en la Figura 22.

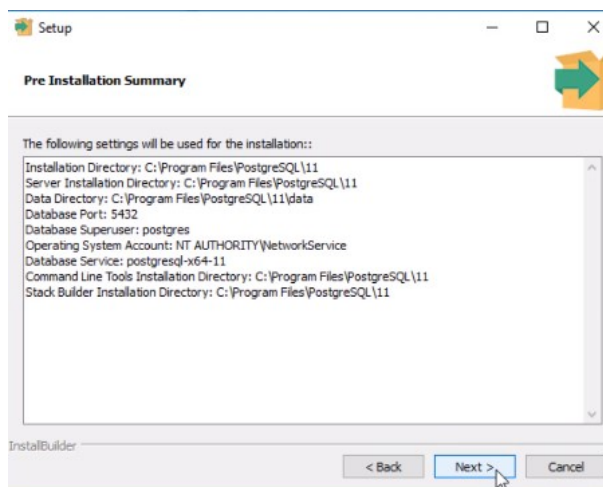


Figura 22: Ventana de verificación de instalación.

IX En la ventana de verificación, dar clic en “Siguiete” para iniciar la instalación, como se muestra en la Figura 23 y Figura 24.

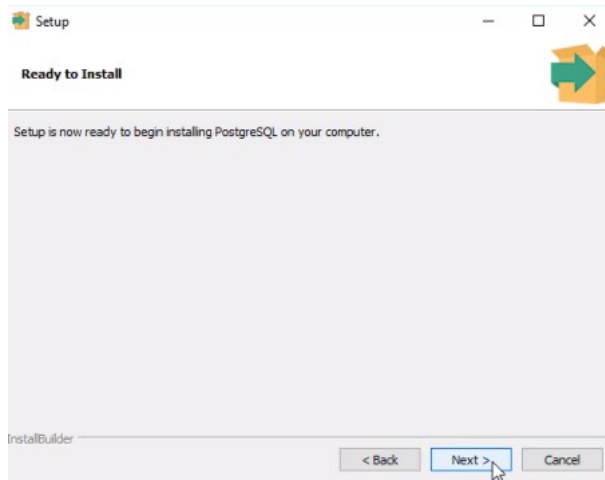


Figura 23: Ventana de verificación e inicio de instalación.

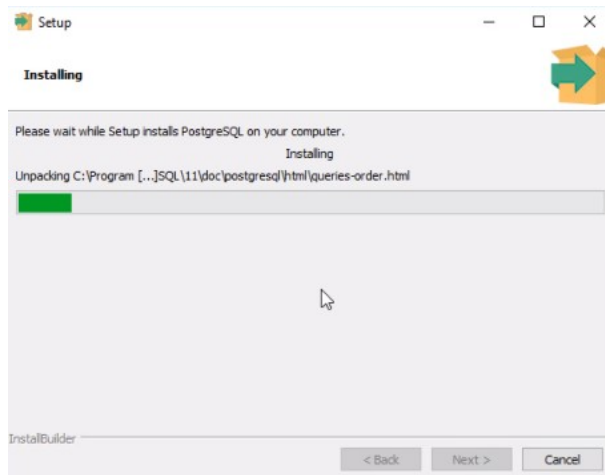


Figura 24: Ventana de avance de instalación.

- x Una vez terminada la instalación, quitar la selección de Stack Builder pues no se necesita. Posteriormente dar clic en “Finalizar”, como se muestra en la Figura 25.



Figura 25: Ventana de selección para instalar Stack Builder.

4. Estudio de caso

Ahora que tienes instalado PostgreSQL es tiempo de cargar tu base de datos. Para este curso contamos con información pública y de libre acceso, esta información fue extraída de los datos de <https://www.imdb.com/>. Las base de datos puede ser encontrada juntos con este material o será proporcionada por el facilitador del curso.

4.1. Especificaciones de la base de datos IMDb

Cada conjunto de datos está contenido en un archivo con formato gzip y valores separados por tabulaciones (TSV) en el juego de caracteres UTF-8. La primera línea de cada archivo contiene encabezados que describen lo que hay en cada columna. Un ‘\N’ se usa para denotar que un campo en particular falta o es nulo para ese título/nombre. Los conjuntos de datos disponibles son los siguientes:

- **title_akas** - Contiene la siguiente información para los títulos:
 - *titleId* - un tconst, un identificador alfanumérico único del título.
 - *ordering* - un número para identificar de manera única las filas de un determinado titleId.
 - *title* - el título localizado.
 - *region* - la región para esta versión del título.
 - *language* - el idioma del título.
 - *types* Conjunto numerado de atributos para este título alternativo. Uno o más de los siguientes: “alternativa”, “dvd”, “festival”, “tv”, “video”, “working”, “origina”, “imdbDisplay”. Es posible que en el futuro se añadan nuevos valores sin previo aviso.
 - *attributes* - Términos adicionales para describir este título alternativo, no enumerados.
 - *isOriginalTitle* - 0: título no original; 1: título original.
- **title_basics** - Contiene la siguiente información para los títulos:
 - *tconst* - identificador alfanumérico único del título.
 - *titleType* - el tipo/formato del título (por ejemplo, película, cortometraje, series de televisión, tvepisode, vídeo, etc.).
 - *primaryTitle* - el título más popular / el título utilizado por los cineastas en los materiales promocionales en el punto de estreno.

- *originalTitle* - título original, en el idioma original.
 - *isAdult* - 0: título no adulto; 1: título adulto.
 - *startYear (AAAA)* - representa el año de lanzamiento de un título. En el caso de las series de televisión, es el año de inicio de la serie.
 - *endYear (AAAA)* - Serie de TV fin de año. N^o para todos los demás tipos de títulos.
 - *runtimeMinutes* - tiempo de ejecución principal del título, en minutos.
 - *genres* - incluye hasta tres géneros asociados al título.
- **title_crew** - Contiene información sobre el director y el escritor de todos los títulos del IMDb. Los campos incluyen:
 - *tconst* - identificador alfanumérico único del título.
 - *directors* - director(es) del título dado.
 - *writers* - escritor(es) del título dado.
- **title_episode** - Contiene la información del episodio de tv. Los campos incluyen:
 - *tconst* - identificador alfanumérico del episodio.
 - *parentTconst* - identificador alfanumérico de la serie de TV matriz.
 - *seasonNumber* - número de temporada al que pertenece el episodio.
 - *episodeNumber* - número de episodio del tconst en la serie de TV.
- **title_principals** - Contiene el reparto/tripulación principal de los títulos:
 - *tconst* - identificador alfanumérico único del título.
 - *ordering* - un número para identificar de manera única las filas de un determinado titleId.
 - *nconst* - identificador alfanumérico único del nombre/persona.
 - *category* - la categoría de trabajo en la que se encontraba la persona.
 - *job* - el título específico del puesto, si procede; si no, "N".
 - *characters* - el nombre del personaje interpretado, si procede, o bien ' \N'.
- **title_ratings** - Contiene la clasificación IMDb y la información de votos para los títulos:
 - *tconst* - identificador alfanumérico único del título.

- *averageRating* - media ponderada de todas las valoraciones individuales de los usuarios.
 - *numVotes* - número de votos que ha recibido el título.
- **name_basics** - Contiene la siguiente información para los nombres:
 - *nconst* - identificador alfanumérico único del nombre/persona.
 - *primaryName* - nombre con el que se acredita más a menudo a la persona.
 - *birthYear* - en formato AAAA.
 - *deathYear* - en formato YYYYYY si procede, si no, ' \N'.
 - *primaryProfession* - las 3 profesiones más importantes de la persona.
 - *knownForTitles* - títulos por los que se conoce a la persona.

4.2. Iniciar PostgreSQL

pgAdmin es una herramienta de código abierto para la administración de bases de datos PostgreSQL. Por lo cual PostgreSQL se ejecutará a través de pgAdmin [12]. PostgreSQL puede ser inicializado de varias formas, la forma más práctica para hacerlo es como se hará en este curso. Para ello, se debe:

1. Ir a la barra de búsqueda de Windows y se debe teclear 'pgAdmin 4' y ejecute el programa, como se muestra en la Figura 26.

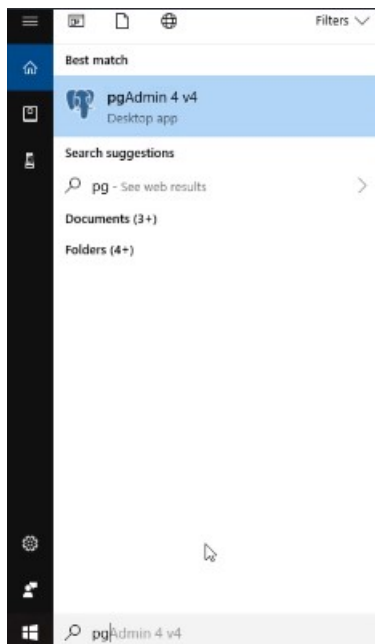


Figura 26: Búsqueda de pgAdmin a través de la barra de inicio de Windows.

NOTA: Las últimas versiones de pgAdmin son corridas en el navegador web predeterminado.

2. Cuando se inicie el programa, del lado izquierdo encontrará el menú. Para cargar un servidor se despliega la lista de servidores en el botón señalado, como se muestra en la Figura 27.

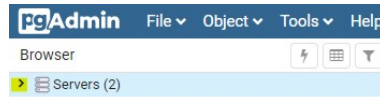


Figura 27: Ventana de selección de servidores dentro del software.

3. Seleccionar PostgreSQL 12 e ingresar la contraseña generada durante la instalación de los softwares. Posteriormente, hacer clic en 'Ok', como se muestra en la Figura 28.

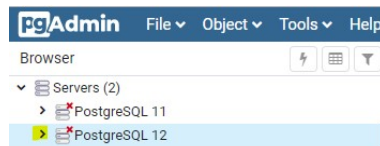


Figura 28: Pestaña de opciones de servidores disponibles.



Figura 29: Ventana de ingreso de contraseña creada durante la instalación

4.3. Iniciar PostgreSQL

1. Dar clic derecho sobre 'Bases de Datos', 'Crear' y posteriormente 'Base de Datos', como se muestra en la Figura 30.

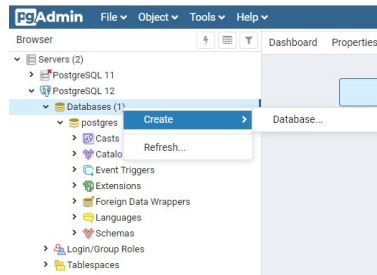


Figura 30: Ventana con las opciones desplegadas para la creación de una nueva base de datos.

2. Nombrar la base de datos 'IMDb' (el nombre es sugerencia, la base de datos puede ser nombrada libremente. Sin embargo, en este documento se referirá como 'IMDb' a la base de datos). Dejar las otras opciones con su configuración predeterminada y dar clic en 'Salvar', como se muestra en la Figura 31.

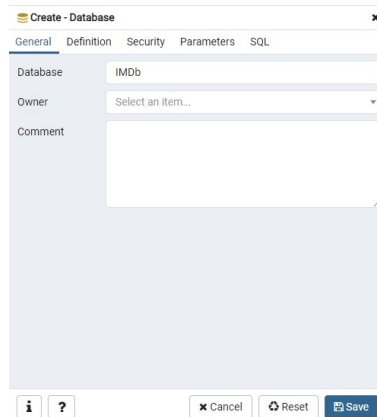


Figura 31: Ventana para nombrar la base de datos.

3. Dar clic derecho sobre la base 'IMDb' y dar clic en 'Restaurar', como se muestra en la Figura 32.

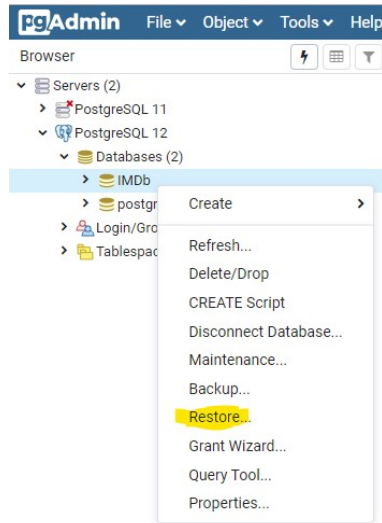


Figura 32: Ventana con las opciones desplegadas para restaurar la base de datos.

4. Buscar la dirección del archivo dando clic en los tres puntos, como se muestra en la Figura 33.

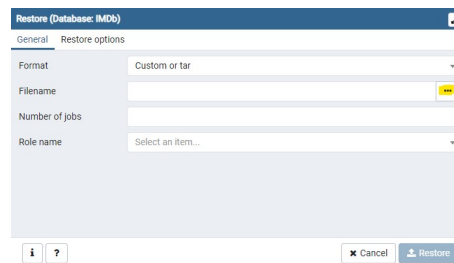


Figura 33: Ventana con las opciones de restauro de la base de datos.

5. Seleccionar el archivo IMDb.tar y dar clic en 'Seleccionar'. Para ello, asegurarse que el formato sea 'Todos los archivos', como se muestra en la Figura 34.

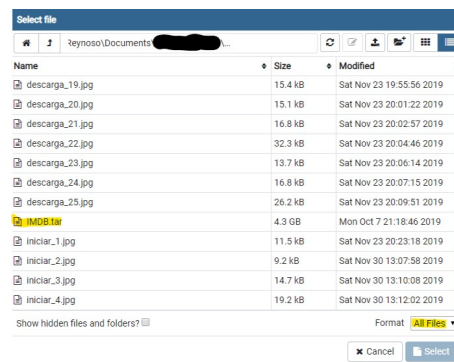


Figura 34: Ventana para cargar el archivo de restauración.

6. Ir a la pestaña de ‘Opciones de Restaurar’. Cambiar a ‘Sí’ las opciones de ‘Post-Data’, ‘Pre-Data’ y ‘Data’, y dejar el resto de las opciones como determinadas. Dar clic en ‘Restaurar’, como se muestra en la Figura 35.

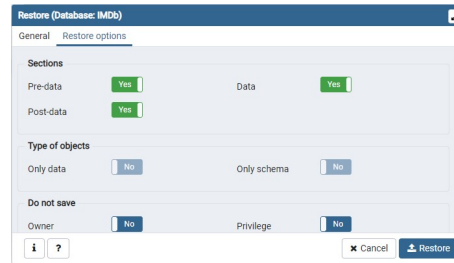


Figura 35: Ventana con las opciones de restaura.

La base de datos se encontrará permanentemente en el programa hasta que sea eliminada manualmente.

5. Instrucciones de extracción de SQL

5.1. Instrucción SELECT (Seleccionar)

Sintaxis de la instrucción SELECT:

```
SELECT columna1, columna2,... FROM nombre_de_tabla
```

- Primero, si se desea especificar una lista de columnas en la tabla de la cual se desea consultar datos en la clausula SELECT. Se usa una coma entre la columnas en caso de que se desee consultar datos de múltiples columnas.
- Si se desea consultar datos de todas las columnas, se puede usar un asterisco (*) como un atajo para todas las columnas.
- Segundo, se indica el nombre de la tabla después de la indicación FROM.
- El lenguaje SQL tiene sensible a mayúsculas y minúsculas. Esto significa que usar SELECT o select es efectivamente igual.
- Por convención, en este curso se usarán los comandos SQL en mayúsculas para hacer el código más fácil de leer.
- No es buena práctica usar el asterisco (*) en la instrucción SELECT. Esto es debido a que, si se desea consultar una tabla grande con varias columnas, la instrucción SELECT

con asterisco (*) consultará todos los datos de las columnas enteras, lo cual podría no ser necesario. Hace que el servidor de la base de datos trabaje más duro e incrementa el tráfico entre el servidor de la base de datos y los programas. Por lo cual, se deberá especificar los nombres de las columnas en la clausula SELECT cada vez que sea posible para obtener sólo los datos necesarios o bien, limitar las líneas con la instrucción LIMIT.

5.1.1. Práctica: Instrucción SELECT

1. Para abrir la herramienta de consulta con lenguaje SQL, de clic derecho sobre la base de datos IMDb y selecciones “Herramienta de consulta”, como se muestra en la Figura 36.

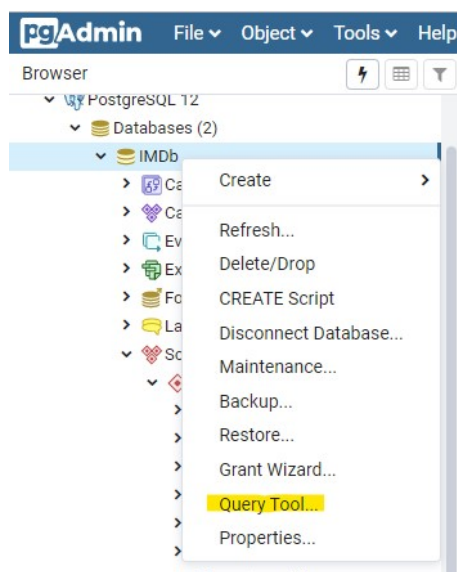


Figura 36: Opciones de acciones de la base de datos, dónde se encuentra la opción para abrir la herramienta de consulta.

La lista de tablas se puede encontrar desplegando IMDb > Schemas > public > Tables.

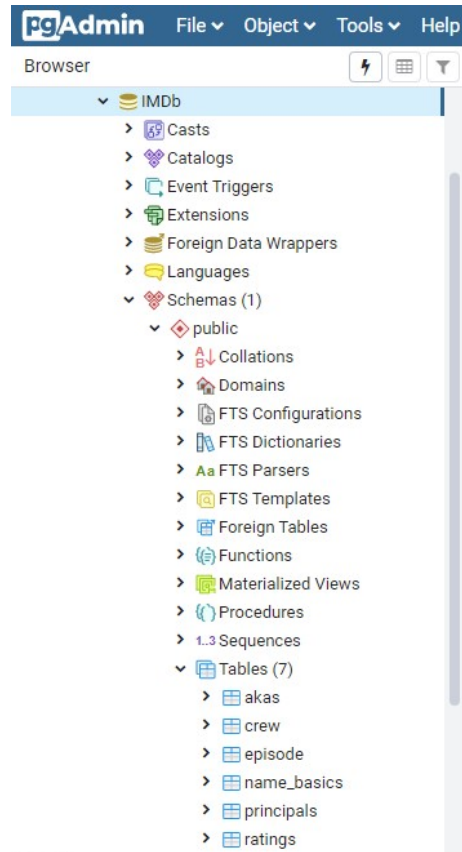


Figura 37: Despliegue de las tablas integradas en la base de datos.

2. Para seleccionar 5 líneas de la tabla “name_basics” Escribir:

```
SELECT
*
FROM name_basics
LIMIT 5
```

Y dar clic en el rayo para correr la instrucción, como se muestra en la Figura 38.

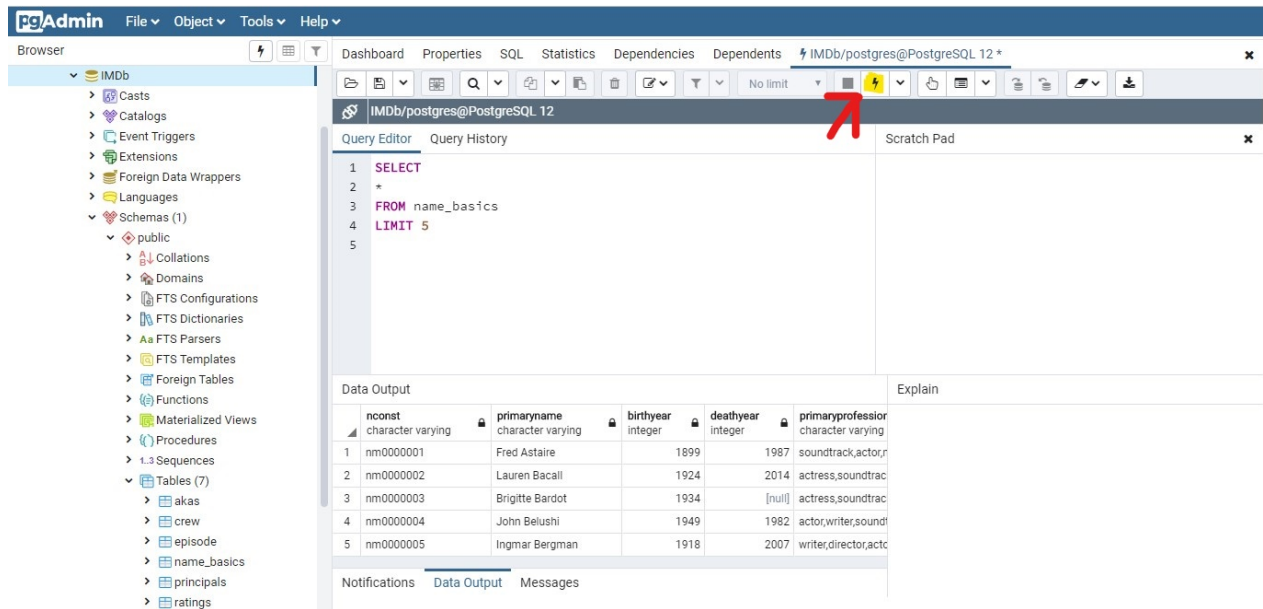


Figura 38: Instrucción de consulta que permite obtener 5 líneas de cualquier tabla.

- Ahora se especificarán las columnas de las cuales deseamos traer información de la misma tabla. Consultaremos los campos “nconst”, “primaryname” y “birthyear”, como se muestra en la Figura 39.

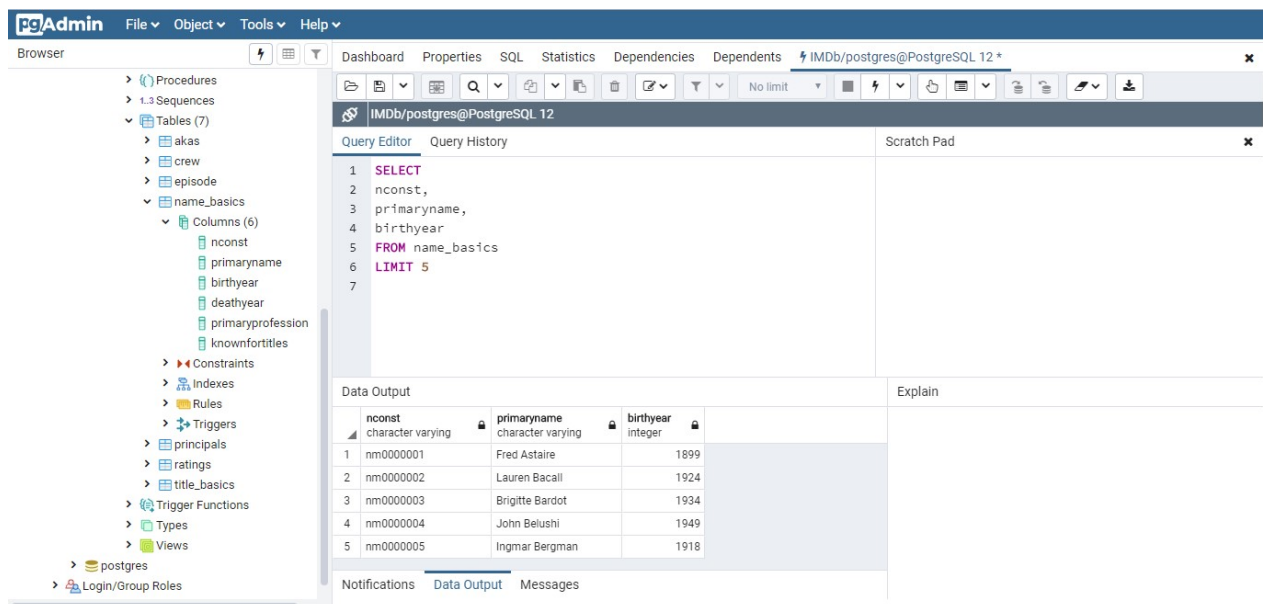


Figura 39: Instrucción con campos específicos que se visualizarán en el resultado.

5.1.2. Ejercicio: Instrucción SELECT

Escribir la instrucción para obtener 7 líneas de todas las columnas de tabla “title_basics”.

Respuesta:

```
SELECT
*
FROM title_basics
LIMIT 7
```

5.2. Instrucción SELECT DISTINCT (Selección Distinta)

- En una tabla, una columna puede contener varios valores duplicados; algunas veces sólo se querrá una lista de valores diferentes (distintos entre sí).
- La instrucción DISTINCT puede ser usada para traer sólo valores distintos.

Sintaxis de la instrucción SELECT DISTINCT:

```
SELECT DISTINCT
columna1,
columna2
FROM nombre_de_tabla
```

5.2.1. Práctica: Instrucción SELECT DISTINCT

Para obtener los distintos tipos de títulos de la tabla “title_basics”.

Respuesta:

```
SELECT
DISTINCT titletype
FROM title_basics
```

Véase solución en Figura 40.

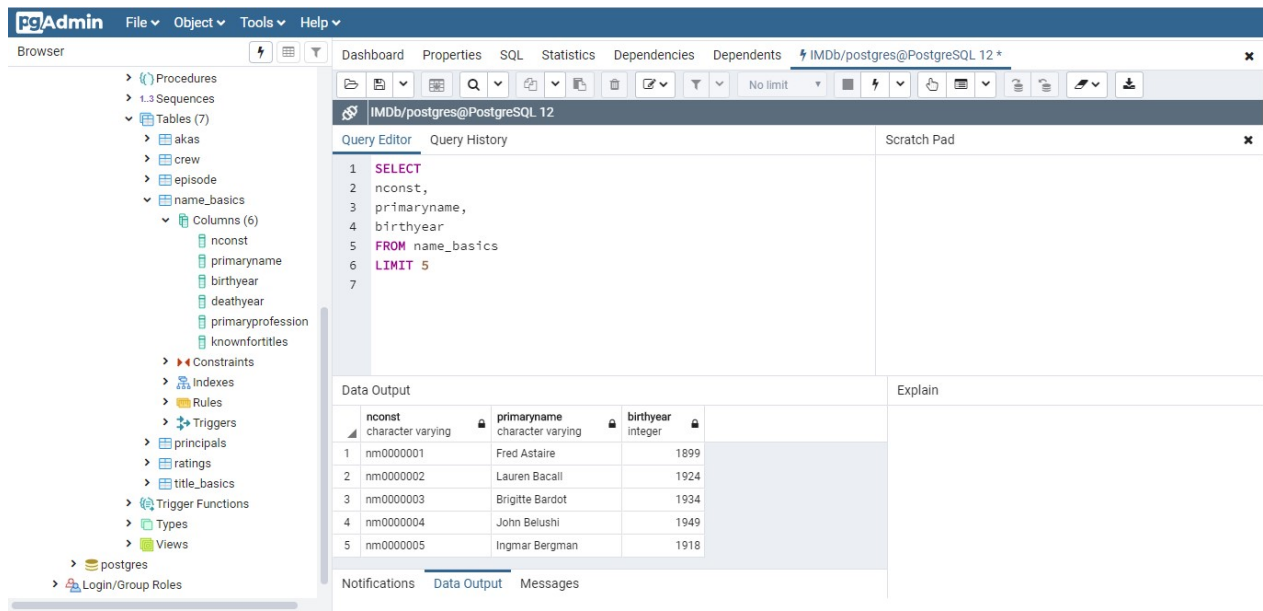


Figura 40: Instrucción SELECT DISTINCT aplicada a los tipos de título.

La instrucción DISTINCT permite que se obtengan valores únicos de la tabla cuando son repetidos. Es decir, la tabla puede contener más de una vez el mismo tipo de título (titletype), al ser posible que varios títulos tengan la misma clasificación, al agregar DISTINCT sólo traerá una sola vez el tipo de título, bastando que esté presente en la tabla una o más veces.

5.3. Instrucción SELECT WHERE (Selección Dónde)

- La cláusula WHERE debe aparecer justo después de la cláusula FROM de la indicación SELECT.
- Las condiciones son usadas para filtrar las filas obtenidas de la indicación SELECT.
- Algunos de los operadores son los presentados en la Figura 41.

OPERATOR	DESCRIPTION
=	Equal
>	Greater than
<	Less than
>=	Greater than or equal
<=	Less than or equal
<> or !=	Not equal
AND	Logical operator AND
OR	Logical operator OR

Figura 41: Operadores de la cláusula WHERE.

5.3.1. Práctica: Instrucción SELECT WHERE

Si se desea seleccionar el nombre principal del actor cuyo nombre es Brad Pitt y su fecha de nacimiento es 1963, el operador lógico AND puede combinar las dos condiciones.

Respuesta:

```
SELECT DISTINCT
primaryname
FROM name_basics
WHERE primaryname = 'Brad Pitt'
AND birthyear = 1963
```

Véase solución en Figura 42.

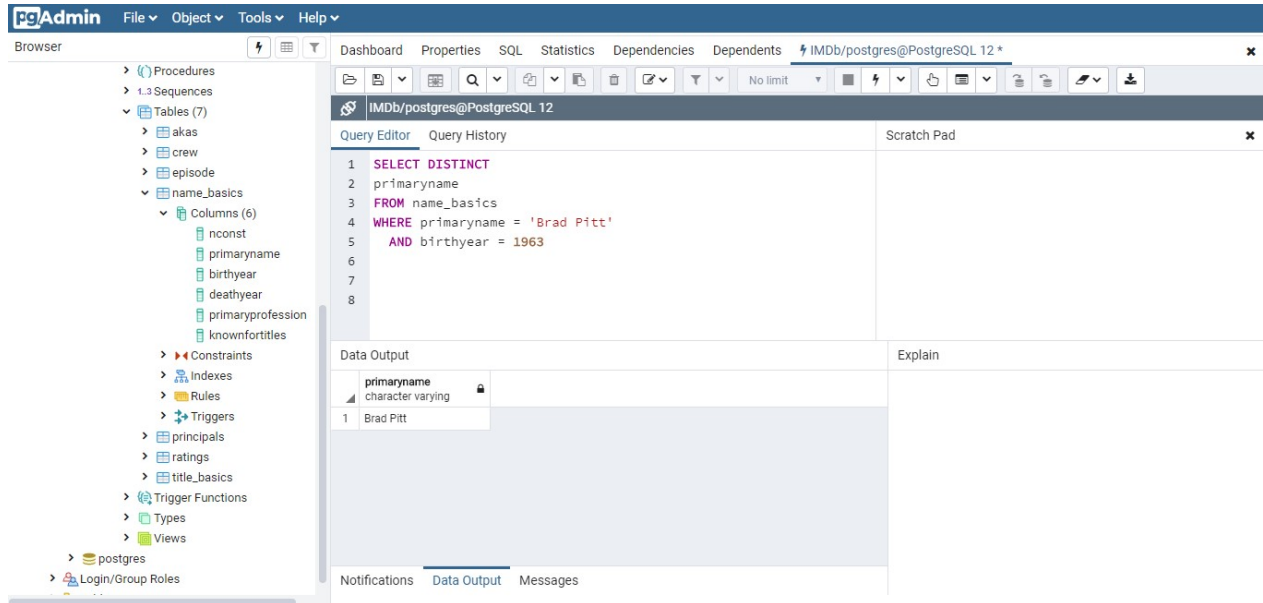


Figura 42: Operadores de la cláusula WHERE aplicados a la tabla name_basics con la función AND.

Nótese que al igual una variable a un texto, el texto debe ir entre comillas simples ('texto'), mientras que al igual a un número se usa una equivalencia simple.

Si se desea traer información del actor Brad Pitt o Johnny Depp, se puede realizar la siguiente consulta con el operador OR.

```
SELECT DISTINCT
```

*

FROM name_basics

WHERE primaryname = 'Brad Pitt' OR primaryname = 'Johnny Depp'

Véase solución en Figura 43.

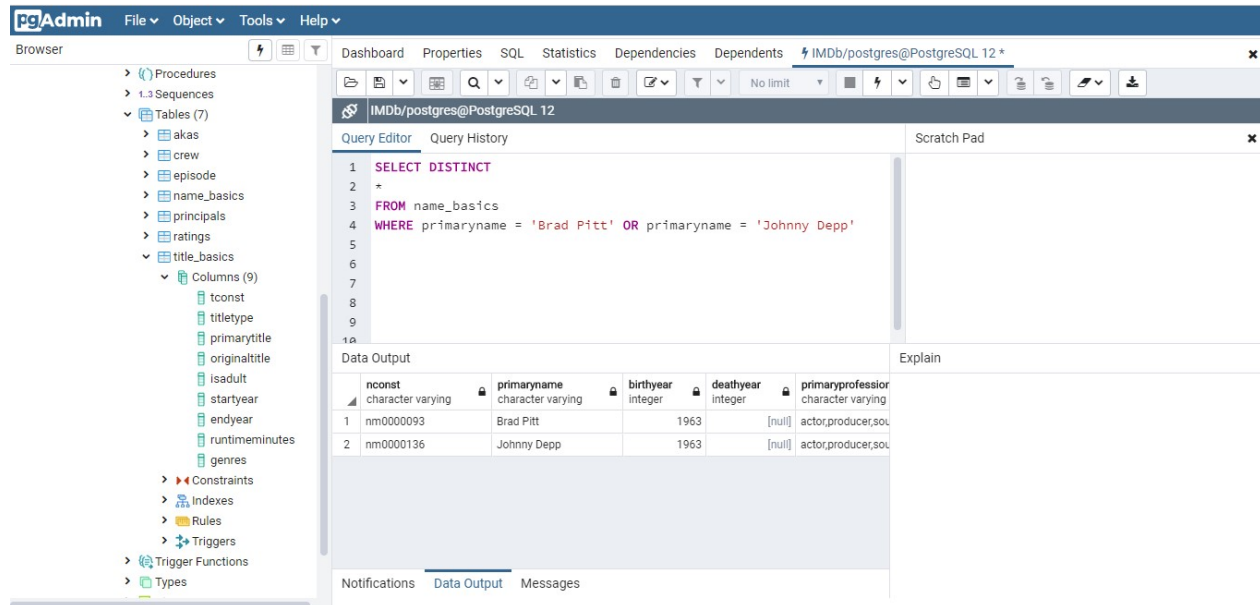


Figura 43: Operadores de la cláusula WHERE aplicados a la tabla name_basics con la función OR.

5.3.2. Ejercicio: Instrucción SELECT WHERE

Escriba la instrucción para obtener todas las columnas de la tabla “title_basics” donde el tipo de título sea 'movie' cuyo año de inicio sea mayor o igual al año 2018.

Respuesta:

```
SELECT
```

```
*
```

```
FROM title_basics
```

```
WHERE titletype = 'movie'
```

```
AND startyear >= 2018
```

6. Instrucciones de limpieza de SQL

6.1. Instrucción COUNT (Contar)

La instrucción COUNT permite obtener el número de filas que coincidan con la condición específica de la consulta.

Sintaxis de la instrucción COUNT:

```
SELECT
COUNT(*)
FROM tabla
```

- La instrucción COUNT(*) obtiene el número de filas obtenidas por la cláusula SELECT.
- Similar a la instrucción COUNT(*), la instrucción COUNT(columna) obtiene el número de filas obtenida en la cláusula SELECT. Sin embargo, no considera los valores nulos en la columna.
- Finalmente, se pueden contar los valores distintos de las columnas, por ejemplo:
SELECT COUNT(DISTINCT columna) FROM table

6.1.1. Práctica: Instrucción COUNT

1. De la table “name_basics” obtener el número de filas de la tabla completa con la siguiente instrucción:

```
SELECT
COUNT(*)
FROM name_basics
```

Véase solución en Figura 44.

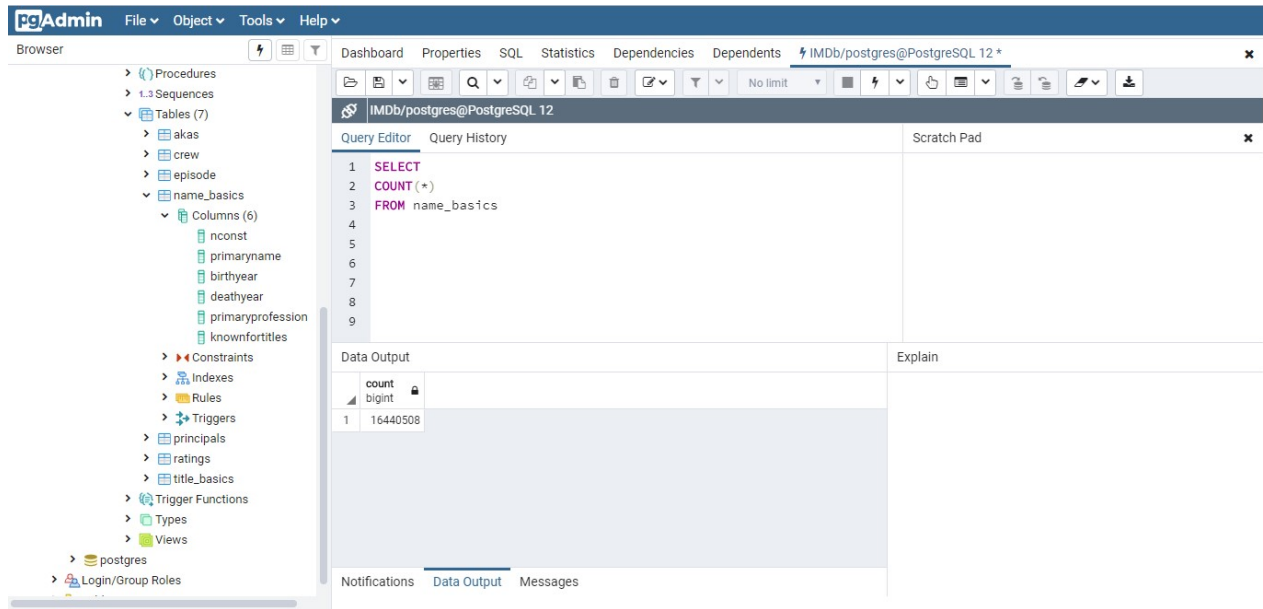


Figura 44: Función COUNT contando todas las líneas de una tabla.

- De la misma tabla, contar el número de años de nacimiento con los que se cuenta registro con la siguiente consulta:

```

SELECT
COUNT(birthyear)
FROM name_basics

```

Véase solución en Figura 45.

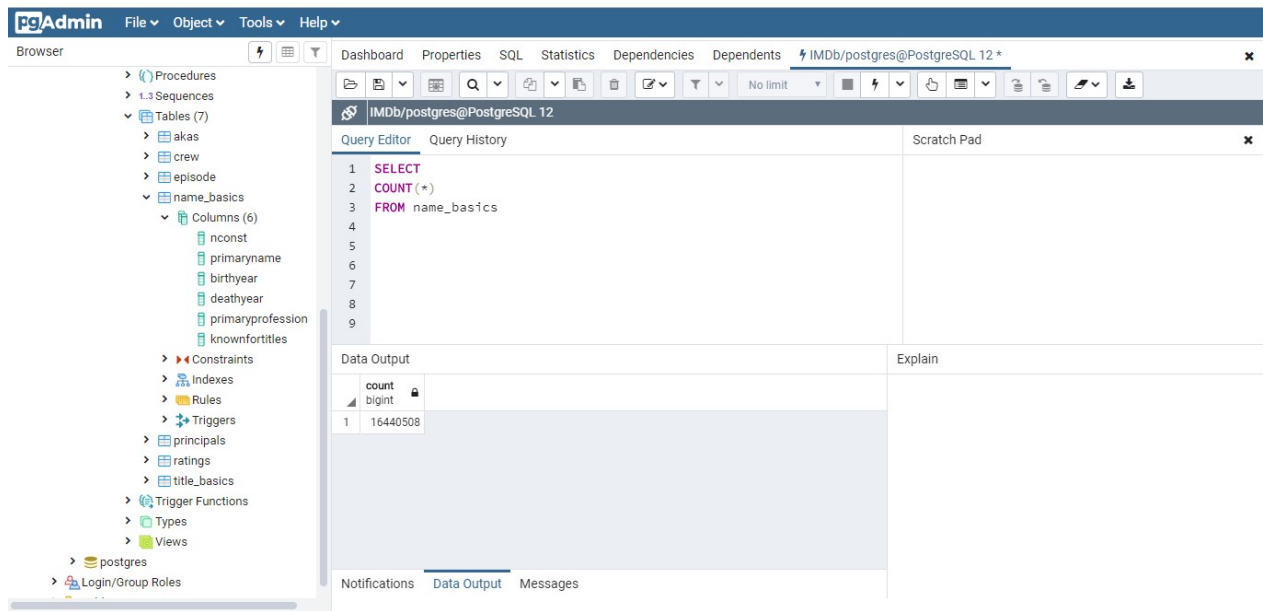


Figura 45: Función COUNT contando todas las líneas de un campo.

Nótese que es menor la cantidad contada, esto es debido a que no todas las líneas cuentan con un registro de años de nacimiento (son nulos).

6.1.2. Ejercicio: Instrucción COUNT

Escriba la instrucción para obtener la cuenta de los tipos distintos de tipo de título de la tabla "title_basics".

Respuesta:

```

SELECT
COUNT(DISTINCT titletype)
FROM title_basics

```

6.2. Instrucción ORDER BY (Order Por)

La instrucción ORDER BY permite ordenar las filas obtenidas de SELECT en orden ascendente o descendente basado en una criterio especificada.

El ejemplo siguiente ilustra la sintaxis de la indicación ORDER BY:

```

SELECT
columna1,
columna2

```

```
FROM nombre_de_tabla
ORDER BY columna1 ASC/DESC
```

- Se especifica la columna por la cual quieres ordenar en la instrucción ORDER BY. Si se desea ordenar los resultados por múltiples columnas, se usa una coma para separar entre columnas.
- Se usa ASC para ordenar los resultados en forma ascendente y DESC para ordenar en forma descendente.
- Si lo deja en blanco, la instrucción ORDER BY usará ASC de facto.
- Los nombres de las columnas pueden ser sustituidos por el número de columnas que ocupan. Ejemplo de tabla “ratings”: ORDER BY tconst, averagerating, nomvotes es igual que ORDER BY 1,2,3.

6.2.1. Práctica: Instrucción ORDER BY

Para obtener los títulos de películas de manera descendente de la tabla “title_basics”, donde obtengamos de la consulta el título original y el año de inicio (es importante eliminar los nulos), escribir la siguiente instrucción:

```
SELECT
originaltitle,
startyear
FROM title_basics
WHERE startyear IS NOT NULL
ORDER BY 2 DESC
```

Véase solución en Figura 46.

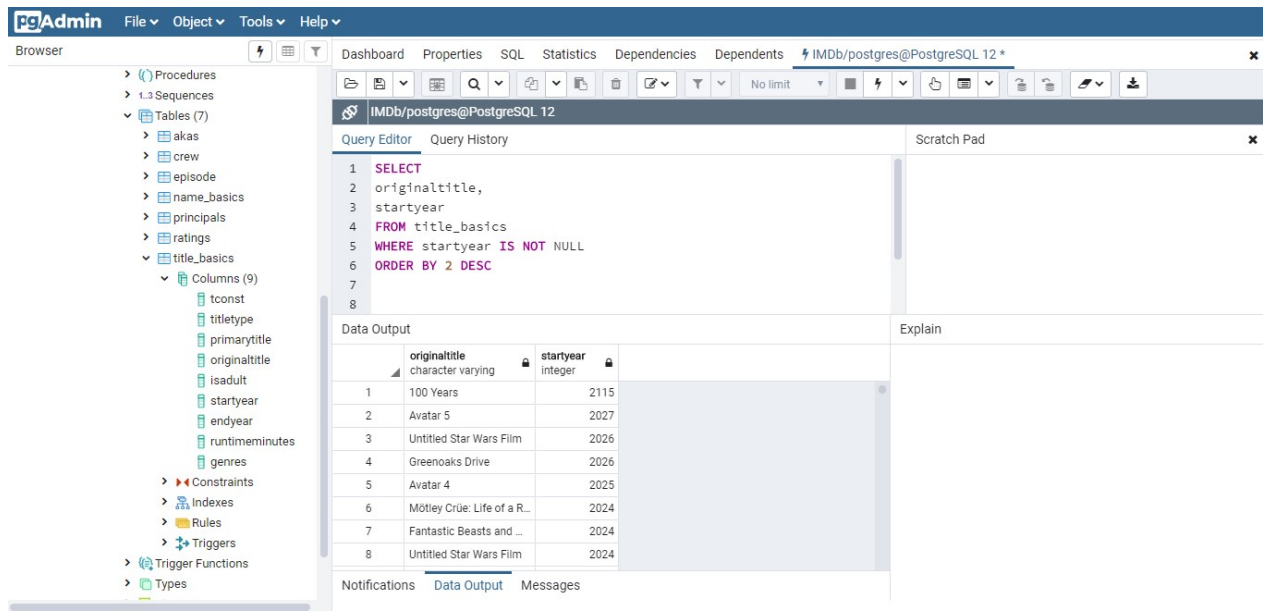


Figura 46: Resultados de la consulta cuyo orden se define como descendente según los valores de la segunda columna.

Se eliminan los valores nulos del año de inicio pues existen muchos registros son año de inicio asignado.

6.3. Instrucción BETWEEN (Entre)

Se usa la instrucción BETWEEN para encontrar valores contra un rango de valores. Por ejemplo:

Valor BETWEEN low AND high

Si se quiere revisar si un valor está fuera de rango, se usa la instrucción NOT BETWEEN de la siguiente manera:

Value NOT BETWEEN low AND high

6.3.1. Práctica: Instrucción BETWEEN

De la tabla “name_basics”, se desea obtener las columnas de nombre primario y años de nacimiento de aquellas personas nacidas entre 1993 y 1995. Se obtiene escribiendo la siguiente instrucción:

```

SELECT
primaryname,
birthyear
FROM name_basics
WHERE birthyear BETWEEN 1993 AND 1995

```

Véase solución en Figura 47.

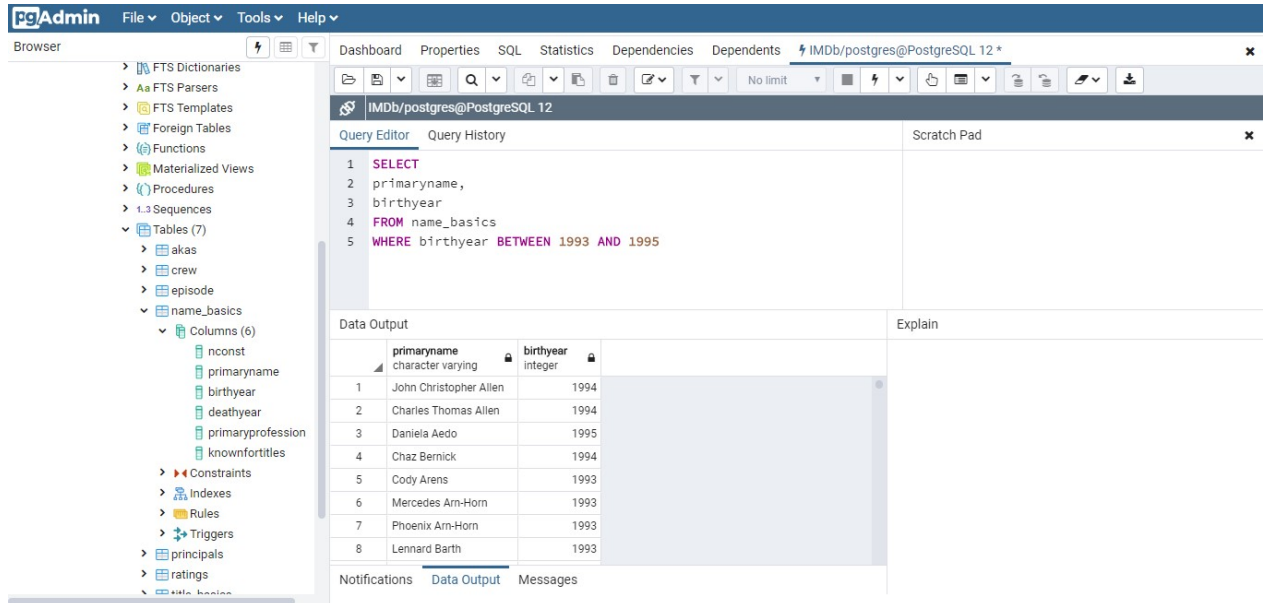


Figura 47: Resultados de consulta de personas nacidas entre 1993 y 1995.

6.4. Instrucción IN (En)

Se usa la instrucción IN con la cláusula WHERE para revisar si algún valor coincide con la lista de valores.

La sintaxis de la operación IN es la siguiente:

variable IN (valor1,valor2,...)

La expresión obtiene verdaderos si algún valor coincide con la lista de valores. La lista de valores no está limitada a una lista de números o textos, sino también como el resultado de una instrucción SELECT como se muestra a continuación:

variable IN (SELECT valor FROM nombre.de.tabla)

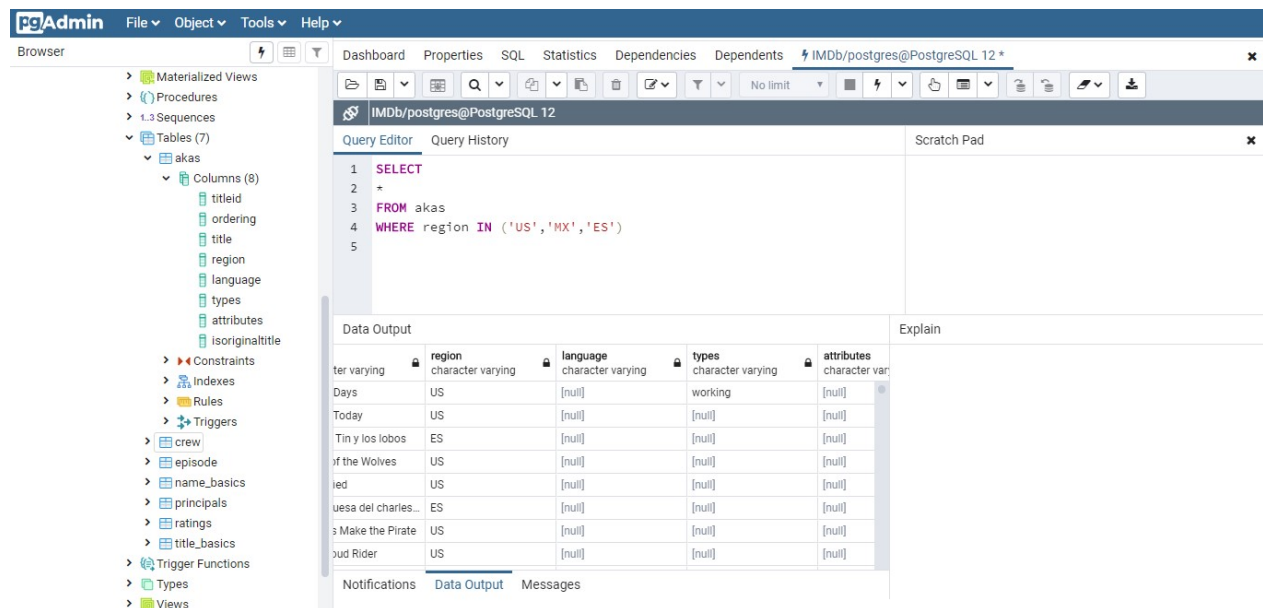
Al igual que en la instrucción BETWEEN, se puede usar NOT para ajustar la instrucción IN (NOT IN).

6.4.1. Práctica: Instrucción IN

De la tabla “akas”, traer toda la información de aquellos títulos desarrollados en las regiones de Estados Unidos (US), México (MX) y España (ES). La información se obtiene con la siguiente consulta:

```
SELECT
*
FROM akas
WHERE region IN ('US','MX','ES')
```

Véase solución en Figura 48.



The screenshot shows the pgAdmin interface with a SQL query editor and a results table. The query is:

```
1 SELECT
2 *
3 FROM akas
4 WHERE region IN ('US','MX','ES')
5
```

The results table has the following columns: region, language, types, and attributes. The data is as follows:

region	language	types	attributes
US	[null]	working	[null]
US	[null]	[null]	[null]
ES	[null]	[null]	[null]
US	[null]	[null]	[null]
US	[null]	[null]	[null]
ES	[null]	[null]	[null]
US	[null]	[null]	[null]

Figura 48: Resultados de consulta de información de películas en tres regiones específicas.

6.5. Instrucción LIKE (Como)

Cuando se desea encontrar un texto del que no recuerda el contenido completo específico. Se puede encontrar dicho texto en la tabla al buscar cualquier valor de la columna como tenga alguna parte del texto que se recuerde.

La sintaxis de la instrucción LIKE se escribe de la siguiente manera:

```
SELECT
columna1,
columna2
FROM nombre_de_tabla
WHERE columna1 LIKE 'texto %'
```

Nótese que en la cláusula WHERE contiene una expresión nueva:

- La columna1, el operador LIKE y una cadena de texto que contiene un signo porcentual (%), al cual se le refiere como “patrón”.
- La consulta trae todas las líneas cuyos valores dentro de la columna columna1 comiencen con el texto seguido de cualquier secuencia de caracteres.
- Esta técnica es conocida como coincidencia por patrones.
- Se construye un patrón al combinar una cadena de texto con carácter comodín y el uso de LIKE o NOT LIKE para encontrar coincidencias:
 - Porcentaje (%) para la coincidencia de cualquier secuencia de caracteres.
 - Barra baja (_) para la coincidencia de caracteres individuales .

6.5.1. Práctica: Instrucción LIKE

Se desea encontrar todas aquellas personas de la tabla “name_basics” cuyo apellido sea Pratt, año de nacimiento y la profesión principal que ejercen. Recordar que en la columna “primaryname” vive el nombre y apellido de la persona. Para realizar la consulta se debe escribir la siguiente consulta:

```
SELECT DISTINCT
primaryname,
primaryprofession,
birthyear
FROM name_basics
WHERE primaryname LIKE '%Pratt'
```

Esto se observa en la Figura 49

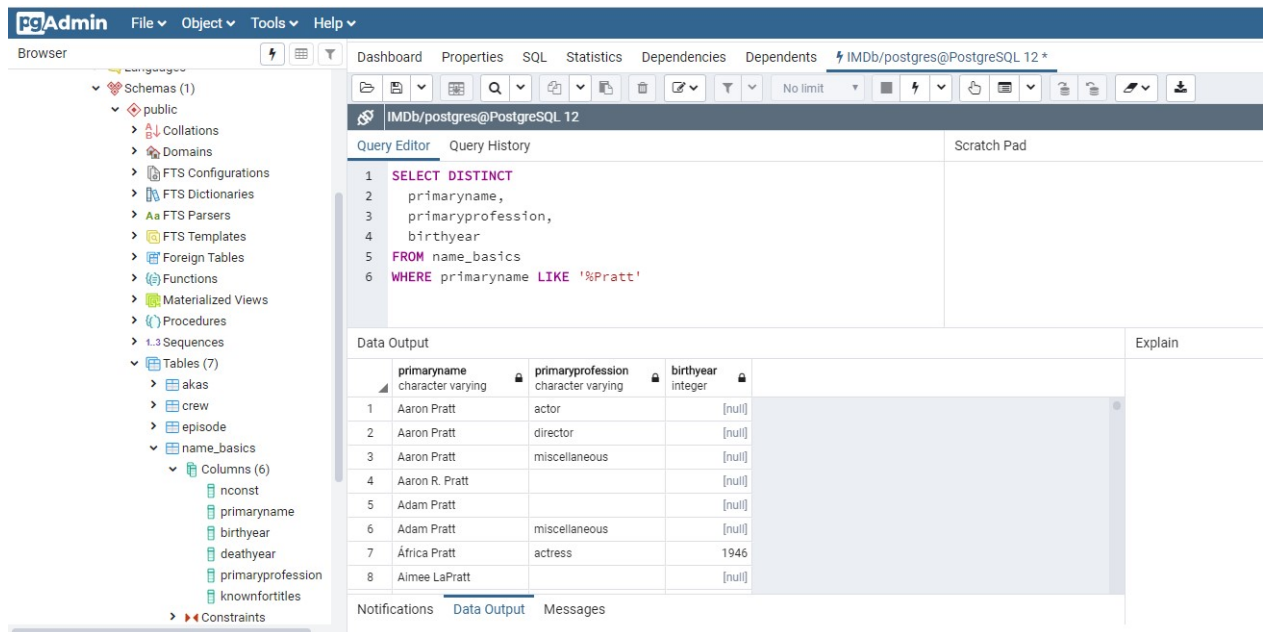


Figura 49: Ejemplificación de consulta para un texto libre del lado izquierdo.

El símbolo de porcentaje puede ser usado en ambos lados del texto, por ejemplo, se quiere encontrar todas aquellas personas que tengan “nn” en alguna parte de su nombre. La consulta sería la siguiente:

```

SELECT DISTINCT
primaryname,
primaryprofession,
birthyear
FROM name_basics
WHERE primaryname LIKE '%nn%'

```

Esto se observa en la Figura 50

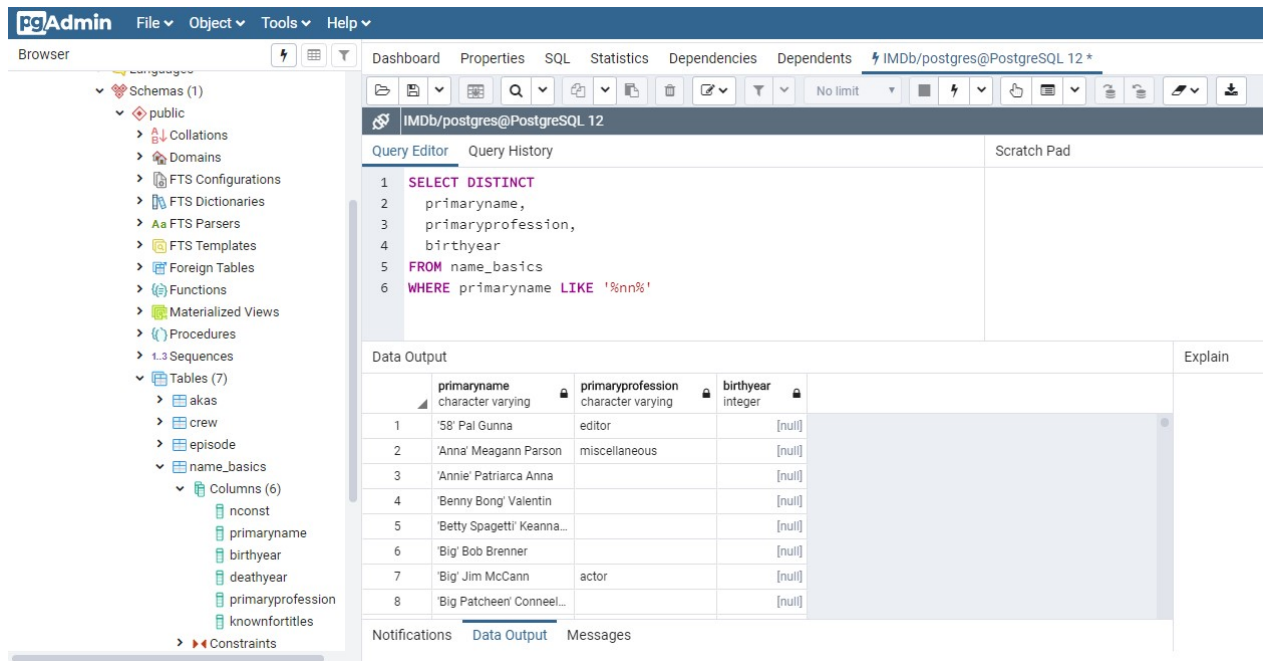


Figura 50: Ejemplificación de consulta para un texto libre de ambos lados.

El signo de guión bajo (_) traer todas aquellas líneas que tengan sólo un carácter a continuación del texto, por ejemplo:

```

SELECT DISTINCT
primaryname,
primaryprofession,
birthyear
FROM name_basics
WHERE primaryname LIKE '_n%'

```

Esto se observa en la Figura 51

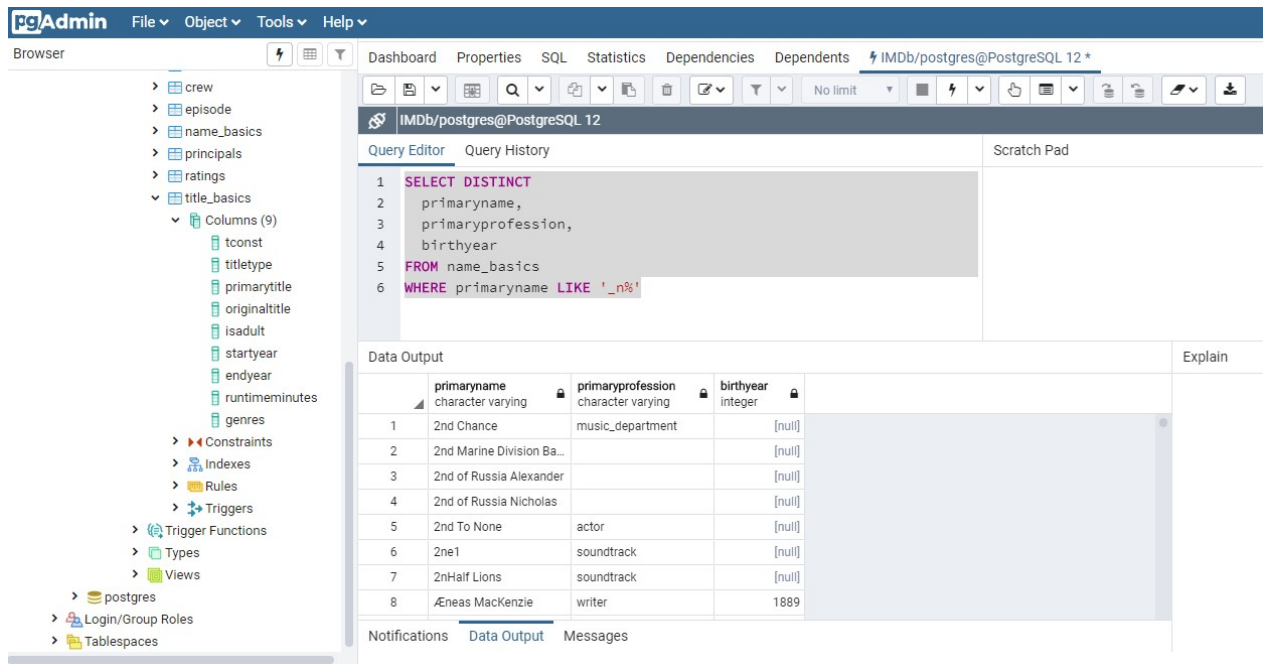


Figura 51: Ejemplificación de consulta para un texto con un sólo caracter del lado izquierdo.

7. Funciones de agregación

Una función de agregación le permite realizar un cálculo sobre un conjunto de valores para devolver un único valor escalar. A menudo se utilizan funciones agregadas con las cláusulas GROUP BY y HAVING de la declaración SELECT.

Las siguientes son las funciones agregadas de SQL más utilizadas:

- **AVG:** calcula el promedio de un conjunto de valores.
- **MIN:** obtiene el valor mínimo en un conjunto de valores.
- **MAX:** obtiene el valor máximo en un conjunto de valores.
- **SUM:** calcula la suma de los valores.

7.0.1. Práctica: Funciones de agregación

De la tabla “title.basics” utilizaremos la columna “startyear” para poder demostrar el uso de MIN y MAX. Primero se traerá la película cuyo año sea el mínimo en la base de datos de la siguiente manera:

```
SELECT
MIN(startyear)
FROM title_basics
```

Esto se observa en la Figura 52.

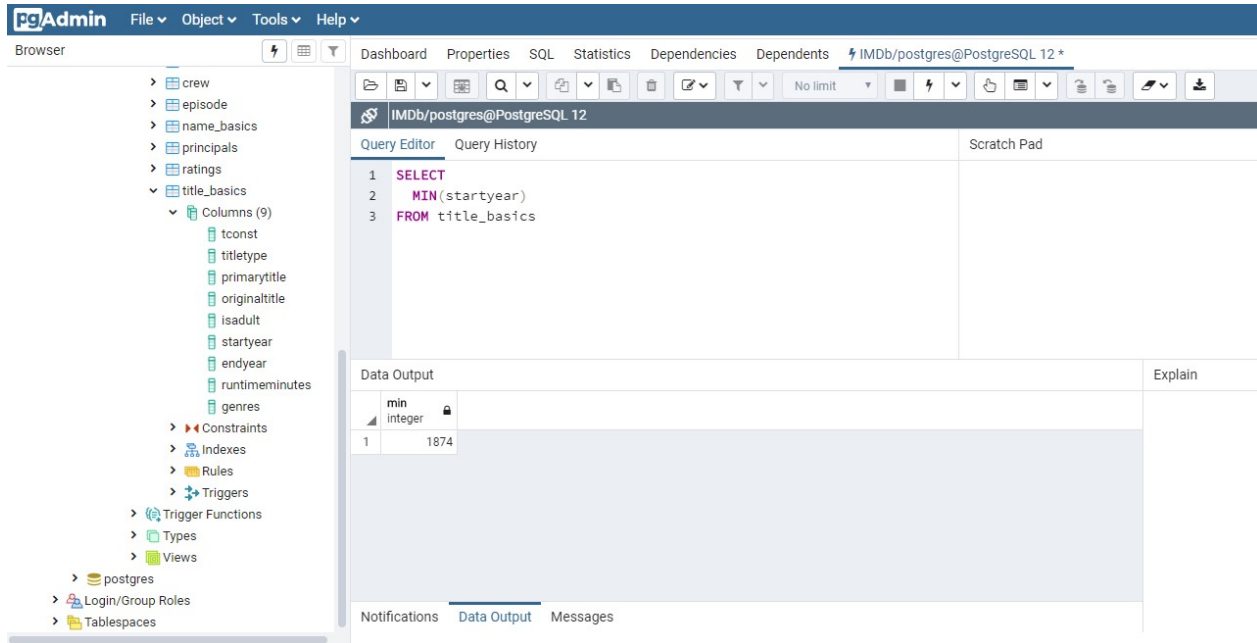


Figura 52: Resultado de la consulta para obtener el año máximo.

Para traer el año que sea el máximo se hace de la siguiente manera:

```
SELECT
MAX(startyear)
FROM title_basics
```

Esto se observa en la Figura 53.

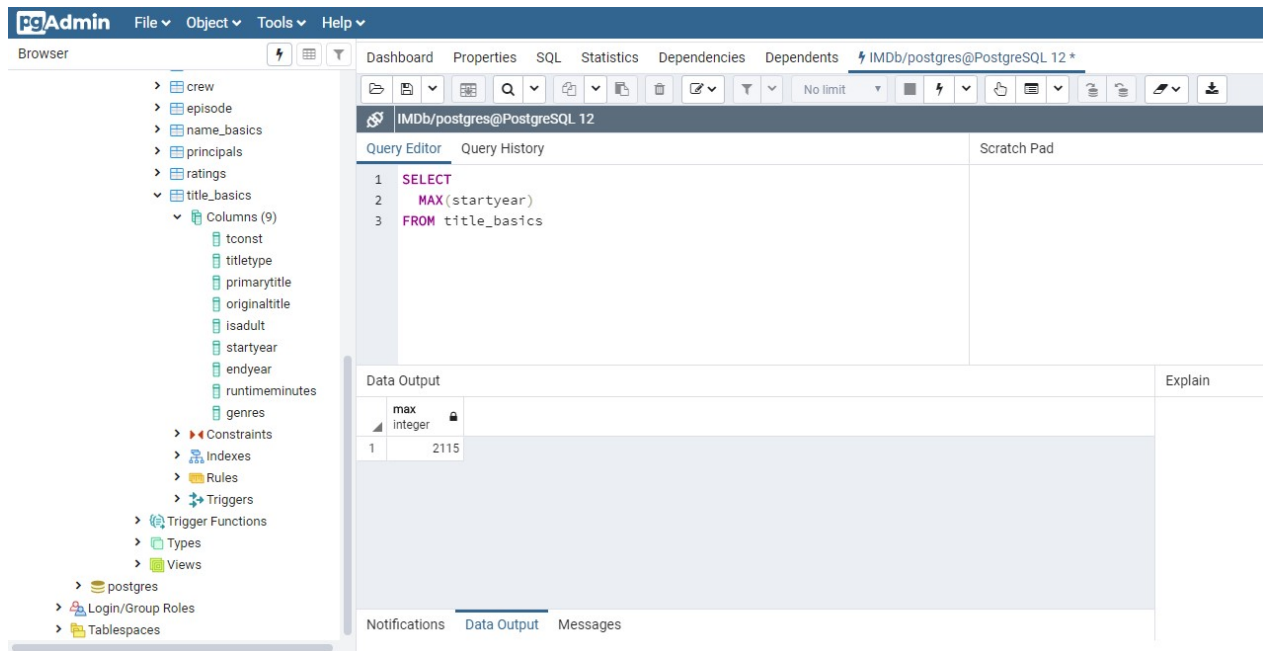


Figura 53: Resultado de la consulta para obtener el número promedio de votos.

Para ver la función AVG obtendremos el promedio de número de votos que existen en la tabla “ratings”. Se obtiene con la siguiente consulta:

```

SELECT
AVG(numvotes)
FROM ratings
  
```

Esto se observa en la Figura 54.

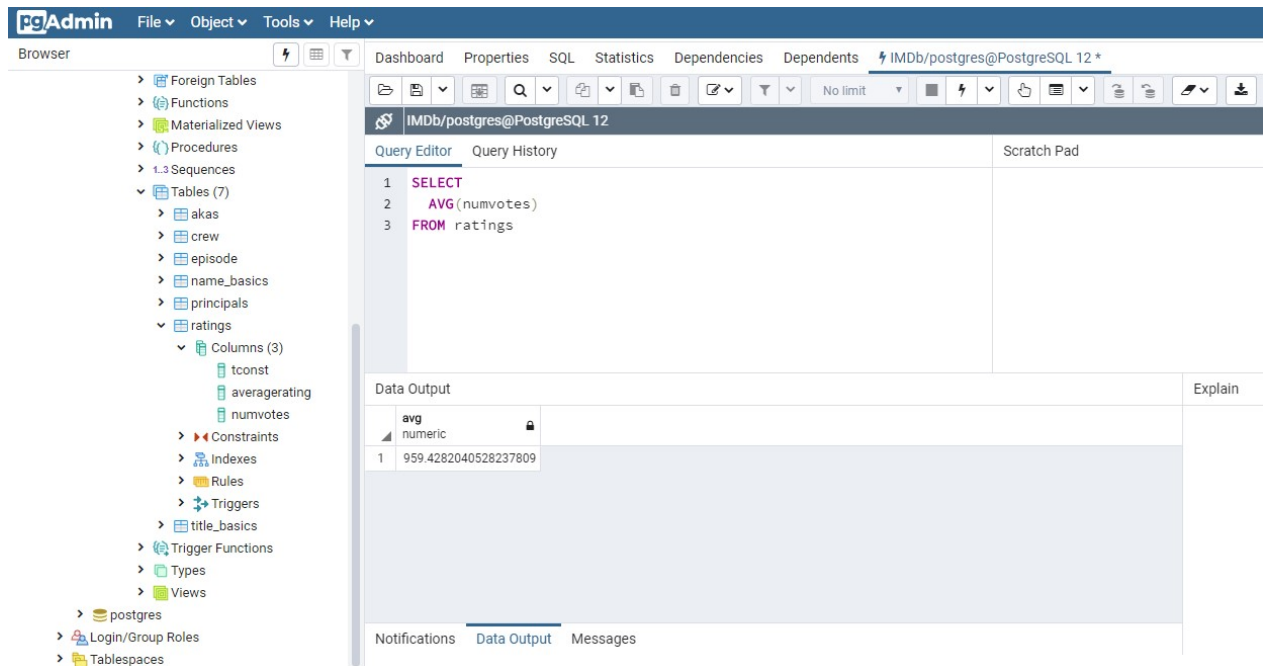


Figura 54: Resultado de la consulta para obtener el número promedio de votos.

Se puede usar la misma tabla y columna para mostrar el uso de SUM. La siguiente consulta suma todos los votos realizados en esta base de datos. Esto se puede observar en la Figura 55.

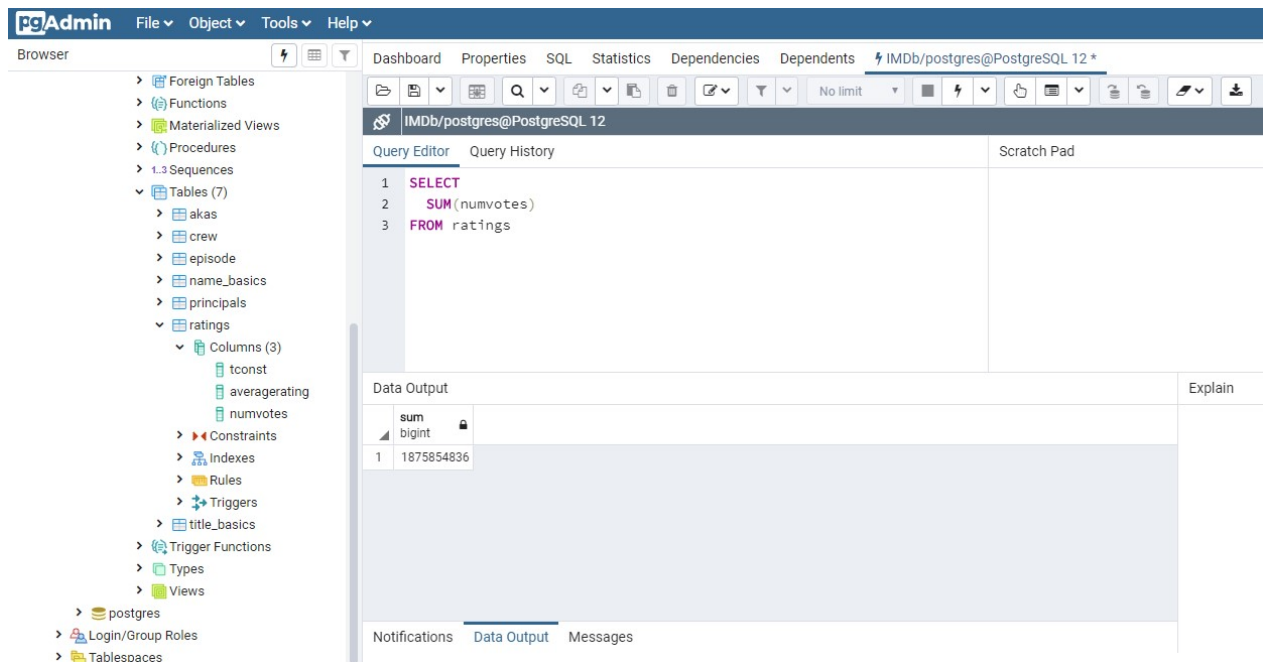


Figura 55: Resultado de la consulta para obtener la suma de los votos.

7.1. Instrucción GROUP BY (Agrupar por)

- La cláusula GROUP BY divide las filas obtenida de la instrucción SELECT en grupos.
- Para cada grupo, se puede agregar una función, por ejemplo:
 - Calcular la suma de los ítems.
 - Contar el número de ítems en grupos.

La sintaxis del GROUP BY es la siguiente:

```
SELECT
columna1,
función_de_agregación(columna2)
FROM tabla
GROUP BY columna1
```

7.1.1. Práctica: Instrucción GROUP BY

Para practicar con la función de agregación COUNT y la instrucción GROUP BY, de la tabla “title_basics” se contarán todos los títulos y se agruparán por tipo de título. Lo anterior se obtiene con la siguiente consulta:

```
SELECT DISTINCT
titletype,
COUNT(*)
FROM title_basics
GROUP BY 1
```

Esto se observa en la Figura 56.

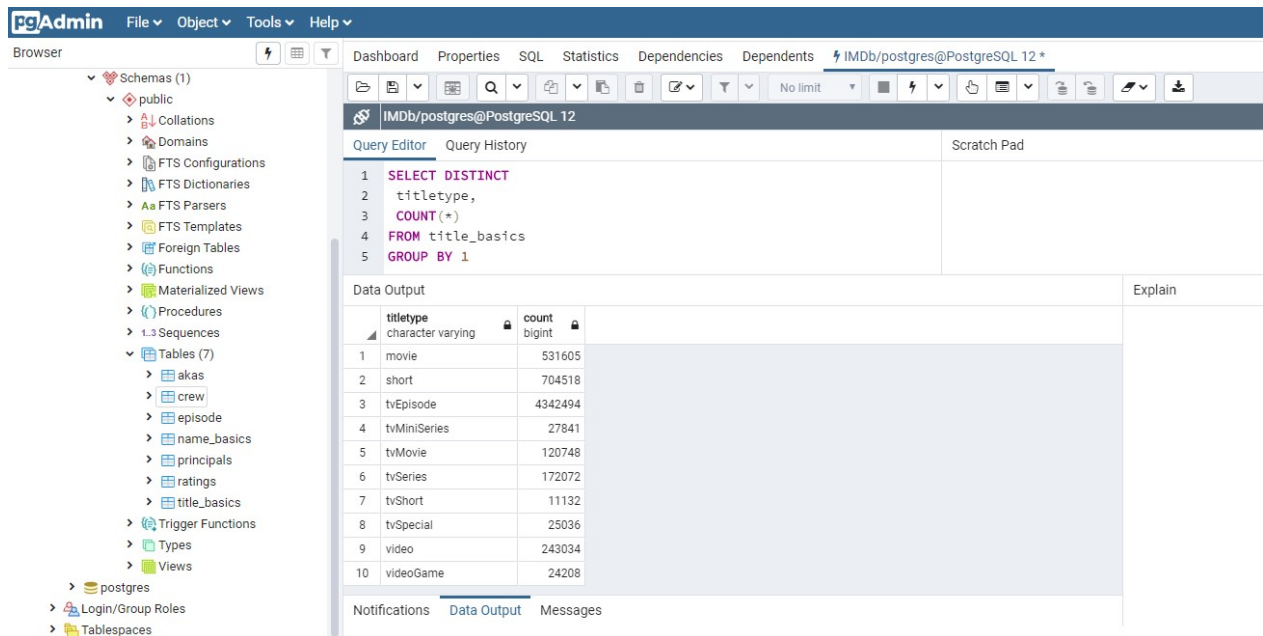


Figura 56: Resultado de la consulta agrupado por el tipo de título.

7.2. Instrucción HAVING (Teniendo)

La instrucción HAVING se usa con en conjunto con GROUP BY para filtrar grupos de filas que no cumplan con una condición específica.

La sintaxis de la instrucción HAVING es la siguiente:

```

SELECT columna1, función.de.agregación(columna2)
FROM tabla
GROUP BY columna1
HAVING condición

```

7.2.1. Diferencia entre HAVING y WHERE

La cláusula HAVING proporciona la condición para un grupo de líneas creado por la cláusula GROUP BY, mientras que la cláusula WHERE proporciona la condición individual a las líneas antes de que la cláusula GROUP BY aplique.

7.2.2. Práctica: Instrucción HAVING

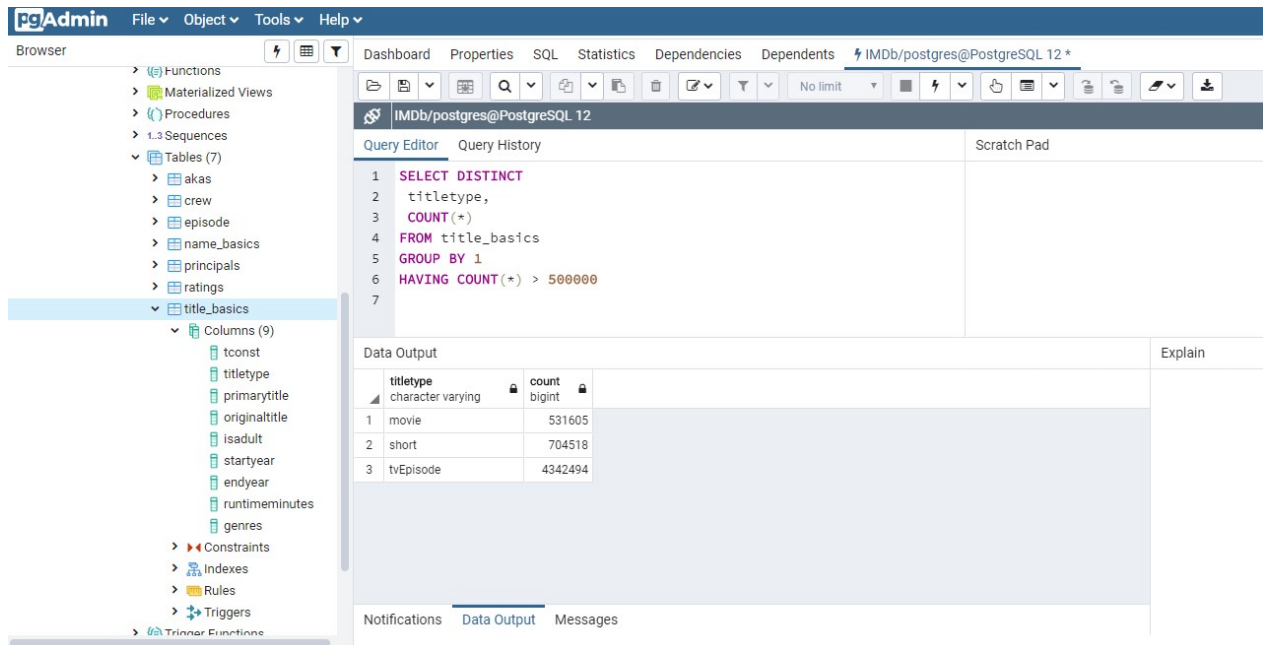
Con el ejemplo de la instrucción anterior, se contarán todos los títulos, se agruparán por tipo de título y se agrega la condición de que sólo traiga aquellos grupos cuya cuenta de títulos sea mayor a 500 000. Lo anterior se obtiene con la siguiente consulta:

```

SELECT DISTINCT
titletype,
COUNT(*)
FROM title_basics
GROUP BY 1
HAVING COUNT(*) > 500000

```

Esto se observa en la Figura 57.



The screenshot shows the PGAdmin interface with a query editor and a data output window. The query editor contains the following SQL code:

```

1 SELECT DISTINCT
2 titletype,
3 COUNT(*)
4 FROM title_basics
5 GROUP BY 1
6 HAVING COUNT(*) > 500000
7

```

The data output window displays the following results:

titletype	count
movie	531605
short	704518
tvEpisode	4342494

Figura 57: Resultado de la consulta de tipos de título que tengan más de 500 mil ejemplares.

8. Instrucción AS (Como)

AS permite renombrar selecciones de columnas o tablas con un alias. Por ejemplo, al tomar una instrucción vista anteriormente como la siguiente:

```

SELECT DISTINCT
titletype,
COUNT(*)
FROM title_basics
GROUP BY 1

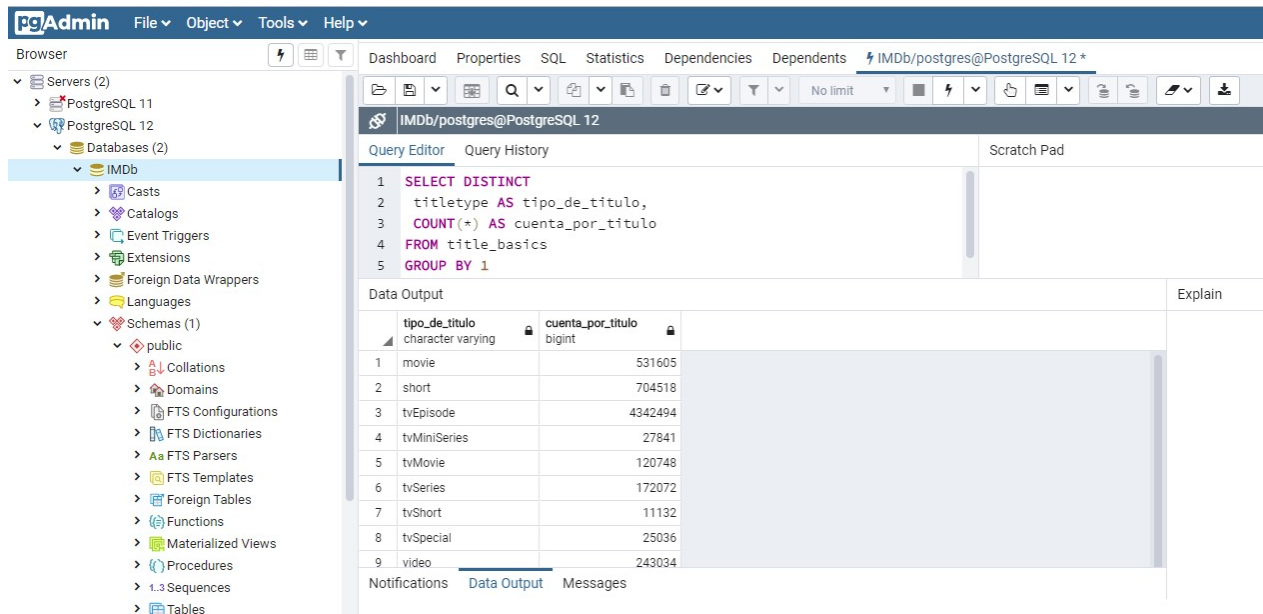
```

Veremos que la primera columna obtiene el mismo título (“titletype”), y la segunda, por ser una función de agregación, obtiene el nombre que le asigna PGAdmin (en este caso

“count”). Ambas columnas pueden ser renombradas con la función AS de la siguiente manera:

```
SELECT DISTINCT
titletype AS tipo_de_titulo,
COUNT(*) AS cuenta_por_titulo
FROM title_basics
GROUP BY 1
```

Esto se observa en la Figura 58.



The screenshot shows the pgAdmin interface with a SQL query editor and a data output table. The query is:

```
1 SELECT DISTINCT
2 titletype AS tipo_de_titulo,
3 COUNT(*) AS cuenta_por_titulo
4 FROM title_basics
5 GROUP BY 1
```

The data output table has the following columns and rows:

	tipo_de_titulo	cuenta_por_titulo
1	movie	531605
2	short	704518
3	tvEpisode	4342494
4	tvMiniSeries	27841
5	tvMovie	120748
6	tvSeries	172072
7	tvShort	11132
8	tvSpecial	25036
9	video	243034

Figura 58: Ejemplo de función AS para nombrar objetos.

9. JOINS (Uniones)

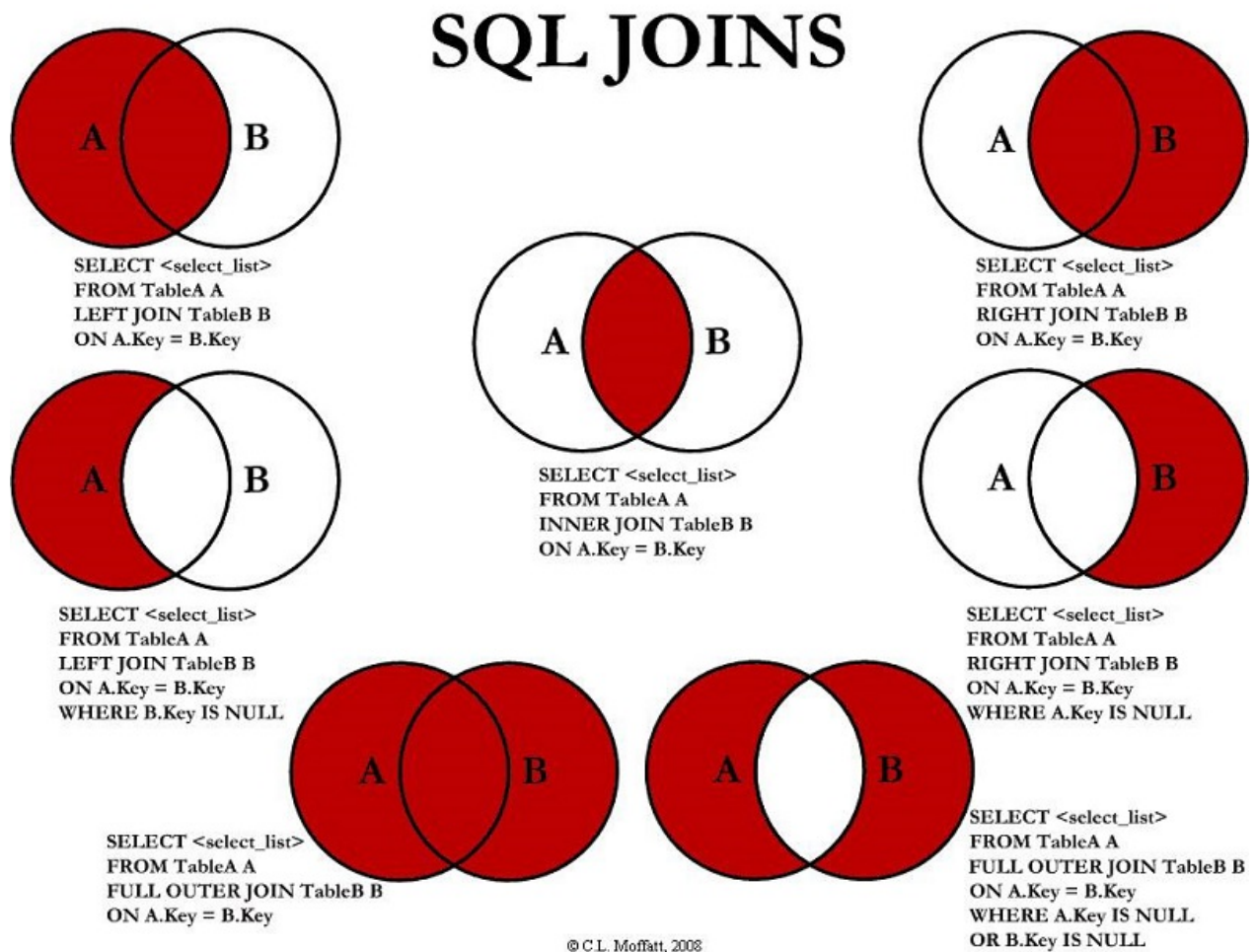


Figura 59: Tipos de JOIN en SQL. Fuente: C.L. Moffatt, 2008.

Hasta ahora se ha aprendido como seleccionar datos de una tabla, escogiendo columnas y filas que se desea e incluso en ordenarlos de manera particular.

Ahora se hablará de uno de los conceptos más importantes en base de datos, estas son las uniones (JOINS). Éstas permiten relacionar datos existentes en una tabla con datos en otras tablas.

Suponga que se desea obtener datos de dos tablas llamadas A y B. La tabla B tiene la llave secundaria que se relaciona con la clave primaria de la tabla A. Esto se ve representado en la Figura 60.

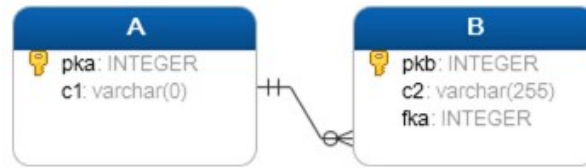


Figura 60: Uniones por medio de llaves primarias. Fuentes: <http://www.postgresqtutorial.com/postgresql-inner-join/>.

Para obtener datos de ambas tablas, se usa la cláusula INNER JOIN (Unión Interna) dentro la indicación SELECT como se muestra a continuación:

```
SELECT
A.clave_primaria,
A.c1,
B.clave_secundaria,
B.c2
FROM A
INNER JOIN B ON A.clave_primaria = B.clave_secundaria
```

- Primero, se especifica la columna en ambas tablas de las que sea desea obtener datos en la cláusula SELECT.
- Segundo, se especifica la tabla principal (e.s. A) en la cláusula FROM
- Tercero, se especifica la tabla con la cual la tabla principal se une (e.s. B) en la cláusula INNER JOIN. Adicionalmente, se agrega la condición por la cual se hará la unión en la palabra clave ON (e.s. A.clave_primaria = B.clave_secundaria)

Para cada fila en la tabla A, se escanea la tabla B para revisar si hay una fila que coincida con la condición (e.s. A.clave_primaria = B.clave_secundaria). Si encuentra una coincidencia, combina las columnas de ambas filas en una sola y agrega la fila combinada en el resultado de la consulta.

Algunas veces las tablas A y B tienen la columna con el mismo nombre, así que debemos especificar a qué columna nos referimos de la siguiente manera:

Nombre_de_tabla.nombre_de_columna

Esto para evitar ambigüedad. Y en caso de que el nombre de la tabla sea largo, se puede usar el alias de la tabla (como se mostró en la instrucción AS).

9.0.1. Práctica: INNER JOIN

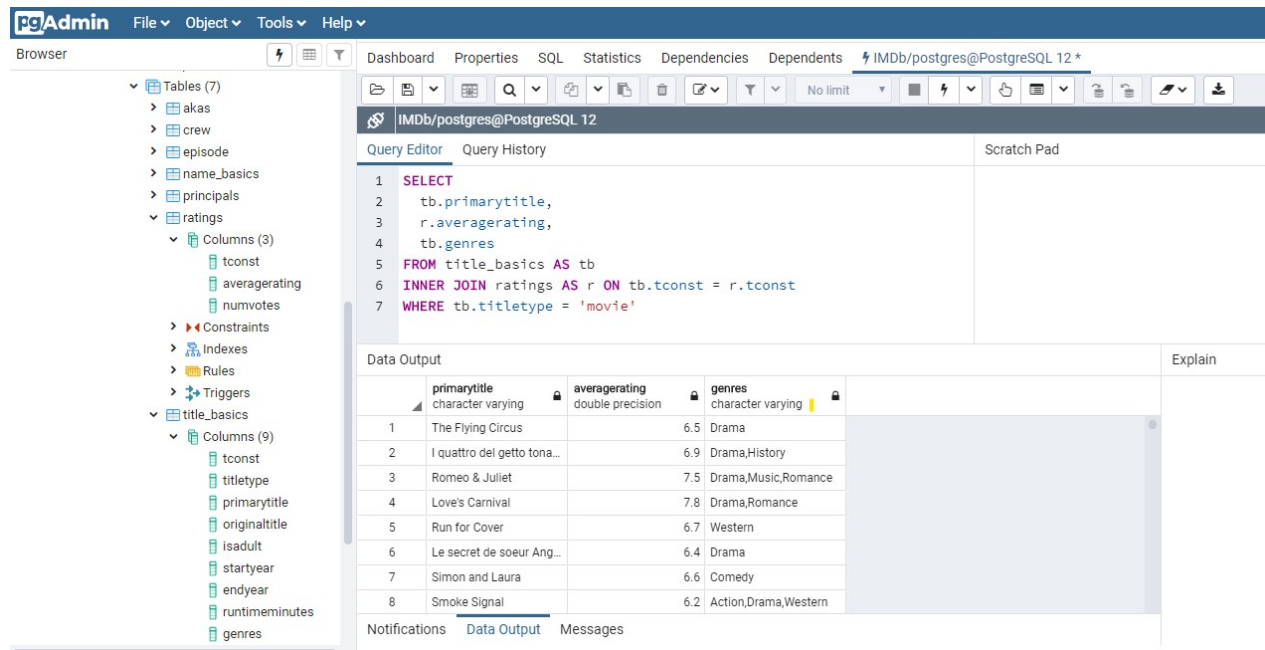
En la base de datos las llaves primarias y secundarias se definen como tconst (clave que se relaciona a los títulos o a las personas, dependiendo de la tabla).

En la práctica uniremos las tablas “title_basics” y “ratings” para crear una tabla con el título primario, el rating promedio y género. Las tablas se unirán por la clave “tconst”, y se filtrarán los tipo de título para sólo obtener aquellos “movie”.

También se practicará la lección anterior renombrando las tablas con la función AS. La tabla “title_basics” como “tb” y “ratings” como “r”. La consulta se realiza de la siguiente manera:

```
SELECT
tb.primarytitle,
r.averagerating,
tb.genres
FROM title_basics AS tb
INNER JOIN ratings AS r ON tb.tconst = r.tconst
WHERE tb.titletype = 'movie'
```

Esto se observa en la Figura 61.



The screenshot shows the pgAdmin interface with a query editor and a data output table. The query editor contains the following SQL code:

```
1 SELECT
2   tb.primarytitle,
3   r.averagerating,
4   tb.genres
5 FROM title_basics AS tb
6 INNER JOIN ratings AS r ON tb.tconst = r.tconst
7 WHERE tb.titletype = 'movie'
```

The data output table shows the following results:

	primarytitle character varying	averagerating double precision	genres character varying
1	The Flying Circus	6.5	Drama
2	I quattro del getto tona...	6.9	Drama,History
3	Romeo & Juliet	7.5	Drama,Music,Romance
4	Love's Carnival	7.8	Drama,Romance
5	Run for Cover	6.7	Western
6	Le secret de soeur Ang...	6.4	Drama
7	Simon and Laura	6.6	Comedy
8	Smoke Signal	6.2	Action,Drama,Western

Figura 61: Unión de dos tablas por medio de una varieable en común.

9.1. Tipos de JOIN

9.1.1. INNER JOIN

Produce sólo el grupo de datos que hacen coincidir ambas tablas [13].

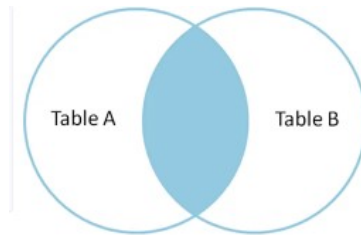


Figura 62: INNER JOIN. Fuente: <http://knowlton-group.com/>

9.1.2. FULL OUTER JOIN

Produce el grupo de datos en la Tabla A y Tabla B, con los datos que coinciden. Si no existe coincidencia, el valor perdido contendrá un nulo [13].

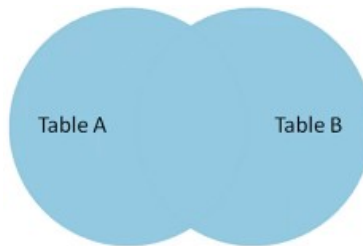


Figura 63: FULL OUTER JOIN. Fuente: <http://knowlton-group.com/>

9.1.3. LEFT OUTER JOIN

Produce un grupo completo de datos de la tabla A, con los valores que coinciden (si están disponibles) en la tabla B. Si no hay coincidencia, el lado derecho contendrá nulos [13].

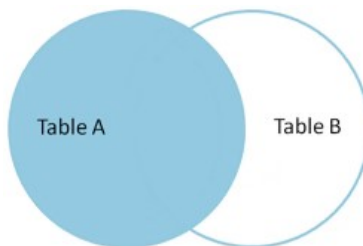


Figura 64: LEFT OUTER JOIN. Fuente: <http://knowlton-group.com/>

9.1.4. LEFT OUTER JOIN con WHERE

Para producir un grupo de datos sólo presentes en la tabla A pero no en la tabla B, se realiza el mismo LEFT OUTER JOIN, después se excluyen los datos no deseados de la tabla del lado derecho con una cláusula WHERE [13].

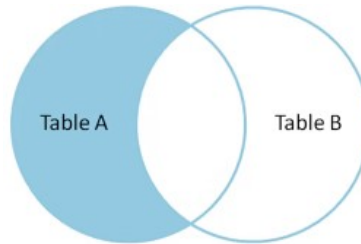


Figura 65: LEFT OUTER JOIN con WHERE. Fuente: <http://knowlton-group.com/>

9.1.5. FULL OUTER JOIN con WHERE

Para producir un grupo de datos únicos entre las tablas A y B, se realiza un FULL OUTER JOIN, después se excluyen los datos no deseados de ambos lados de las tablas con un cláusula WHERE [13].

10. Sub-consultas

Se desean obtener los títulos tipo “movie” cuyo año de inicio sea igual o mayor a 1960, estén por arriba del promedio de calificación de las películas bajo las mismas condiciones y la región sea Estados Unidos (“US”). En la consulta debemos obtener el título original, la calificación promedio por película y el año de inicio.

Ahora se puede traer la información de la siguiente manera:

```
SELECT
tb.originaltitle,
r.averagerating,
tb.startyear
FROM titlebasics AS tb
INNER JOIN ratings AS r ON r.tconst = tb.tconst
INNER JOIN akas AS ak ON ak.tconst = tb.tconst
WHERE tb.startyear >= 1960
AND tb.titletype = 'movie'
AND ak.region = 'US'
```

Sin embargo, aún no estamos trayendo los títulos por encima o igual al promedio de las películas bajo las mismas condiciones. Para poder lograr lo anterior se debe realizar una subconsulta. Una sub-consulta es una consulta anidada dentro otra consulta.

Una sub-consulta es una consulta anidada dentro otra consulta.

Para construir una sub-consulta, se coloca la segunda consulta en paréntesis y se usa en la cláusula WHERE como una expresión:

```
SELECT
tb.originaltitle,
r.averagerating,
tb.startyear
FROM title_basics AS tb
INNER JOIN ratings AS r ON r.tconst = tb.tconst
INNER JOIN akas AS ak ON ak.titleid = tb.tconst
WHERE tb.startyear >= 1960
AND tb.titletype = 'movie'
AND ak.region = 'US'
AND r.averagerating >= (SELECT AVG(r.averagerating)
FROM ratings AS r
INNER JOIN title_basics AS tb ON r.tconst = tb.tconst
WHERE tb.startyear >= 1960
AND tb.titletype = 'movie'
AND ak.region = 'US')
```

11. Sub-tablas

Suponga que se desea obtener el promedio de películas que ha realizado un solo director. Para esto se usará como ejemplo la tabla “crew” como tabla principal, de esta tabla no podemos obtener el promedio de películas realizadas pues todas las variables son texto. Sin embargo, se puede contar el número de títulos por director con la cláusula SELECT DISTINCT y la función COUNT:

```
SELECT
directors,
COUNT(cr.tconst) AS num_movies
FROM crew
```

De la consulta anterior obtenemos un valor numérico en la nueva variable creada llamada “num_movies”. Para practicar lo visto anteriormente crearemos restricciones y uniremos tablas, dichas restricciones serán que los títulos sólo sean del tipo “movie”, del año 1960 en adelante (de la tabla “title_basics”) y que el idioma sea inglés (“en”) (de la tabla akas). La consulta sería la siguiente:

```
SELECT
directors,
COUNT(cr.tconst) AS num_movies
FROM crew AS cr
INNER JOIN title_basics AS tb ON tb.tconst = cr.tconst
INNER JOIN akas AS ak ON ak.titleid = cr.tconst
WHERE tb.titletype = 'movie'
AND tb.startyear >= 1960
AND ak.language = 'en'
GROUP BY 1
```

Aún no se obtiene el promedio de películas por director. Para ello se necesita una sub-tabla.

Las sub-tablas son consultas que crean tablas temporales para las consultas. Es decir, crea tablas que tienen las mismas características de una tabla y pueden ser consultadas bajo las mismas reglas.

12. Ejercicios de repaso

12.1. Instrucción SELECT

Escribir la instrucción para obtener 7 líneas de las columnas “tconst”, “originaltitle” y “startyear” de la tabla “title_basics”.

Respuesta:

```
SELECT
tconst,
originaltitle,
startyear
FROM title_basics
LIMIT 7
```

12.2. Instrucción SELECT DISTINCT

Escribir la instrucción para obtener la distinta profesión principal de la tabla “name_basics”.

Respuesta:

```
SELECT DISTINCT
primaryprofession
FROM name_basics
```

12.3. Instrucción SELECT WHERE

Escriba la instrucción para obtener todas las columnas de la tabla “title_basics” donde el tipo de título no sea ‘movie’ cuyo año de inicio sea mayor o igual al año 2018.

Respuesta:

```
SELECT
*
FROM title_basics
WHERE titletype != 'movie'
AND startyear >= 2018
```

12.4. Instrucción COUNT

Escriba la instrucción para obtener la cuenta de las líneas cuyo tipo de título sea 'videoGame'.

Respuesta:

```
SELECT
COUNT(*) FROM title_basics
WHERE titletype = 'videoGame'
```

12.5. Instrucción ORDER BY

Escribir la instrucción para obtener en orden descendente los promedios de ratings, y los números de votos de la tabla "ratings", se debe eliminar el promedio de votos nulos.

Respuesta:

```
SELECT
*
FROM ratings
WHERE averagerating IS NOT NULL
ORDER BY
averagerating DESC,
numvotes DESC
```

12.6. Instrucción BETWEEN

Escribir la instrucción para obtener de la tabla "title_basics" el título original, el título primario y el año de estreno de todos aquellos títulos de tipo 'movie' que se hayan estrenado entre los años 2018 y 2020.

Respuesta:

```
SELECT
originaltitle,
primarytitle,
startyear
FROM title_basics WHERE startyear BETWEEN 2018 AND 2020
AND titletype = 'movie'
```

12.7. Instrucción IN

Escribir la consulta que traiga los títulos primarios de la tabla “title_basics” de aquellos títulos realizados en la región de México (MX). TIP: Para realizar esta instrucción IN, se debe escribir una subinstrucción SELECT de la tabla “akas”. La variable “tconst” de la tabla “title_basics” y “titleid” de la tabla “akas” son llaves iguales.

Respuesta:

```
SELECT
primarytitle
FROM title_basics
WHERE tconst IN (SELECT titleid FROM akas WHERE region = 'MX')
```

12.8. Instrucción LIKE

Traer de la tabla “title_basics” el título original, año de inicio, tiempo de duración en minutos, género de todas las películas cuya primera palabra de su título sea “Avengers” y tipo de título sea movie.

Respuesta:

```
SELECT DISTINCT
originaltitle,
startyear,
runtimeminutes,
genres
FROM title_basics
WHERE primarytitle LIKE '%Avengers%'
AND titletype = 'movie'
```

12.9. Instrucción GROUP BY

Obtener el número de títulos cuyo tipo de título es ‘movie’ y agruparlo por años de manera descendente eliminando todos aquellos títulos que no tienen un año asignado.

Tip: Para eliminar los títulos con años nulos, usar la instrucción IS NOT NULL.

Respuesta:

```
SELECT DISTINCT
startyear,
COUNT(*) FROM title_basics
```

```
WHERE titletype = 'movie'  
AND startyear IS NOT NULL  
GROUP BY 1  
ORDER BY 1 DESC
```

12.10. Instrucción HAVING

Obtener el número de títulos cuyo tipo de título es 'movie', agruparlo por años de manera descendente eliminando todos aquellos títulos que no tienen un año asignado y filtrar las filas para obtener sólo los años con más de 15000.

Tip: Para eliminar los títulos con años nulos, usar la instrucción IS NOT NULL.

Respuesta:

```
SELECT DISTINCT  
startyear,  
COUNT(*)  
FROM title_basics  
WHERE titletype = 'movie'  
AND startyear IS NOT NULL  
GROUP BY 1  
HAVING COUNT(*) > 15000  
ORDER BY 1 DESC
```

12.11. Instrucción INNER JOIN

Hacer unión de tres tablas, "title_basics" como principal, "akas" y "ratings" como secundarias. Es resultado debe tener las columnas título primario (de "title_basics"), título (de "akas"), región, idioma y calificación promedio de todos los tipos de título "movie" cuya región sea "MX".

Respuesta:

```
SELECT  
tb.primarytitle,  
ak.title,  
ak.region,  
ak.language,  
r.averagerating  
FROM title_basics AS tb
```

```
INNER JOIN akas AS ak ON tb.tconst = ak.titleid
INNER JOIN ratings AS r ON tb.tconst = r.tconst
WHERE tb.titletype = 'movie'
AND ak.region = 'MX'
```

12.12. Sub-consulta

Obtener el título original, la calificación promedio y el año de inicio de los títulos tipo “movie”, de México (“MX”), del año 1990 en adelante y que tenga arriba del promedio del número de votos de las películas con las mismas características.

Respuesta:

```
SELECT
tb.originaltitle,
r.averagerating,
tb.startyear
FROM title_basics AS tb
INNER JOIN ratings AS r ON r.tconst = tb.tconst
INNER JOIN akas AS ak ON ak.titleid = tb.tconst
WHERE tb.startyear >= 1990
AND tb.titletype = 'movie'
AND ak.region = MX
AND r.numvotes >= (SELECT AVG(r.numvotes)
FROM ratings AS r
INNER JOIN title_basics AS tb ON r.tconst = tb.tconst
WHERE tb.startyear >= 1990
AND tb.titletype = 'movie'
AND ak.region = MX)
```

13. Conclusiones

Los datos se han convertido en el bien máspreciado de las compañías, sin embargo, existen una cantidad limitada de ingenieros industriales capaces de extraerlos. En estos apuntes se trabaja con ejercicios prácticos que ejemplifican problemas típicos de manipulación de datos en las organizaciones. Los conocimientos que pretenden estos apuntes son la base para la exploración de bases de datos y generar una mayor familiarización con el lenguaje SQL. A pesar de ello, está limitado a conocer las consultas básicas de datos y sus funciones de agregación. Es tarea del ingeniero tomar estos conocimientos, complementarios a su formación para llevarlos a un análisis verdadero que pueda ser útil para él y la organización. Esto ayuda a que la formación de un ingeniero sea más integral y que desee adentrarse en la Inteligencia de Negocios.

Referencias

- [1] David L. (Junio 20, 2015). Constructivism. Abril 29, 2018, de Learning Theories Sitio web: <https://www.learning-theories.com/constructivism.html>
- [2] David L. (Junio 19, 2015). Cognitivism. Abril 29, 2018, de Learning Theories Sitio web: <https://www.learning-theories.com/cognitivism.html>
- [3] David L. (Enero 31, 2007). Behaviorism. Abril 29, 2018, de Learning Theories Sitio web: <https://www.learning-theories.com/behaviorism.html>
- [4] Gesvin Romero. (Septiembre 27, 2017). 9 Teorías de Aprendizaje más Influyentes. Abril 29, 2019, de Desafíos Educativos del Siglo XXI Sitio web: <https://educar21.com/inicio/teorias-de-aprendizaje-mas-influyentes/>
- [5] Georgetown University. (2018). Teaching as a Process. Abril 29, 2019, de Center for New Designs in Learning & Scholarships Sitio web: <https://cndls.georgetown.edu/atprogram/twl/teaching-as-process/>
- [6] Martin G. Brooks, Jacqueline Grennon Brooks. (1999). The Courage to Be Constructivist. Abril 29, 2019, de Association for Supervision and Curriculum Development Sitio web: <http://www.ascd.org/publications/educational-leadership/nov99/vol57/num03/The-Courage-to-Be-Constructivist.aspx>
- [7] Alice Thomas (Octubre 26, 2010). Understanding The Learning Process To Effectively Differentiate Instruction. Abril 30, 2019, de The Center of Development and Learning Sitio web: <https://www.cdl.org/articles/understanding-the-learning-process-to-effectively-differentiate-instruction/>
- [8] Coordinación de Sistemas - DGPL (2019). Portal de Estadística Universitaria. Septiembre 30, 2019, de Universidad Nacional Autónoma de México Sitio web: <http://www.estadistica.unam.mx/perfiles/>
- [9] WM. A. Wulf, George M. C. Fisher. (Marzo 21, 2002). A Makeover for Engineering Education. Abril 30, 2019, de Issues in Science and Technology Sitio web: https://issues.org/p_wulf/
- [10] Depto. de Ingeniería Industrial (2016). Programas Académicos de Ingeniería Industrial. Abril 30, 2019, de Facultad de Ingeniería, UNAM Sitio web: http://www.ingenieria.unam.mx/programas_academicos/licenciatura/industrial.php
- [11] Interactive Online SQL Training (2020). What is SQL? Julio 30, 2019, de SQL Course Sitio web: <http://www.sqlcourse.com/intro.html>

- [12] Anónimo (Noviembre, 2017). PgAdmin. Febrero 17, 2020, de V-Espino Stio web: <http://www.v-espino.com/~chema/daw1/tutoriales/postgres/pgadmin1.htm>
- [13] Anónimo (Noviembre 26, 2019). SQL Server Joins. Febrero 17, 2020, de SQL Server Sitio web: <https://www.sqlservertutorial.net/sql-server-basics/sql-server-joins/>