

Analisando dados abertos semiestruturados com banco convergente

Daniel S. Panizzo Agosto 2021







Este trabalho está licenciado sob uma Licença Creative Commons Atribuição-Compartilhalgual 4.0 Internacional. Para ver uma cópia desta licença, visite http://creativecommons.org/licenses/by-sa/4.0/.

Guia para Laboratório Hands-On

Introdução	4
Lab 1. Acessando seu ambiente	5
Lab 2. Criando uma instância de Autonomous Data Warehouse	9
Iniciando a criação da Instância de Autonomous Data Warehouse	10
Lab 3. Explorando o Autonomous Data Warehouse	15
Overview do Autonomous Data Warehouse	
Acessando os detalhes do Autonomous Data Warehouse	17
Overview da Service Console do Autonomous Data Warehouse	19
Executando queries no Autonomous Data Warehouse	23
Carregando dados no Autonomous Data Warehouse	25
Lab 4. Consultando dados semi-estruturados no banco convergente .	28
Retornando dados em formato tabular	29
Retornando dados em formato JSON	30
Retornando dados em formato XML	31

Introdução

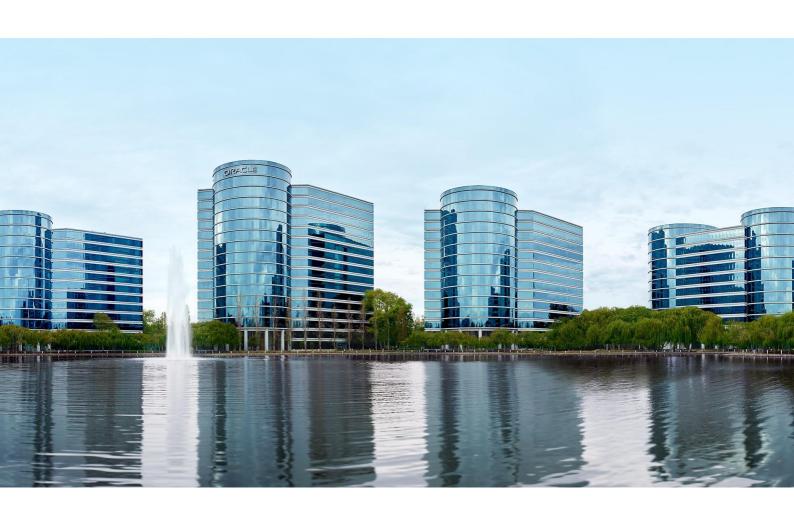
Neste laboratório prático, vamos trabalhar na criação de uma instância de Autonomous Data Warehouse seguindo processos e boas práticas de implementação, assim como criar e exectar as instruções SQL para explorar os multiplos modelos de um banco de dados convergente.

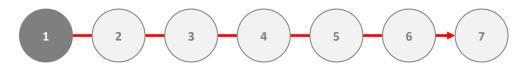
Exploraremos alguns dos recursos disponíveis no Autonomous Data Warehouse:

• SQL Developer Web: Versão Web do conhecido SQL Developer da Oracle.

É importante que os conceitos fundamentais desses recursos estejam claros para uma boa experiência em nossa nuvem.

Lab 1. Acessando seu ambiente





Lab 1. Acessando seu ambiente

Objetivos

- Acessar o console da Oracle Cloud
- Conhecer os serviços de infraestrutura e plataforma
- Familiarizar-se com o ambiente

Nesta seção você aprenderá mais sobre o acesso inicial ao ambiente.

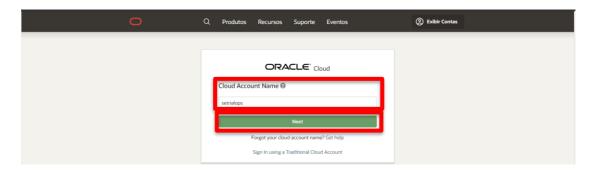
Vá para <u>oracle.com</u>. Você pode alterar o idioma dessa página antes do acesso ao ambiente:



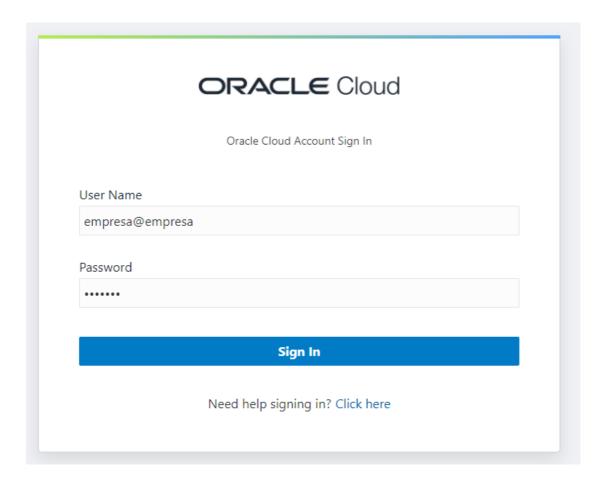
No site já em português, clique em Exibir Contas e depois em Faça Login na Nuvem:



O login deve ser feito com o "Cloud Account Name", onde somente é necessário informar o **nome da conta** (definido no momento de solicitação do trial ou do ambiente final).



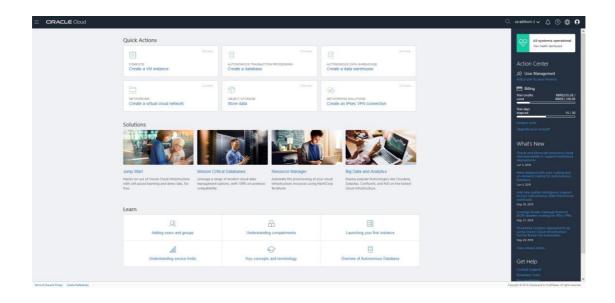
A tela de login para usuário é apresentada. O usuário administrador é identificado pelo e-mail utilizado no cadastro do ambiente.



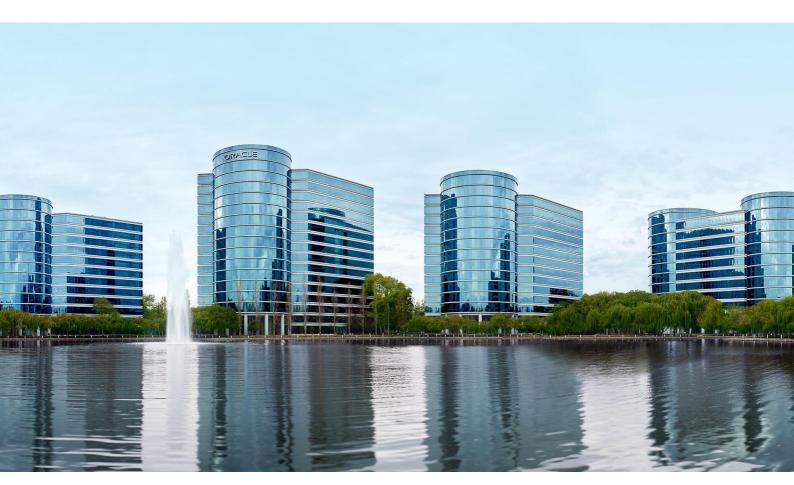
No primeiro acesso é solicitado que sua senha seja alterada.

A tela principal do seu ambiente é apresentada. Nela, você consegue ver algumas ações rápidas para a criação de alguns recursos, alguns artigos de soluções dentro da nuvem da Oracle que possa ajudar, a parte de Learn que leva para a documentação, que é muito bem detalhada.

Na barra superior tem-se o menu que lista todas as abas da console da nuvem, a lupa para pesquisas no ambiente de nuvem, informação de qual a região que está sendo visualizada, no caso da imagem abaixo está sendo visualizado a região de Ashburn, o sino é aonde é feito os anúncios relacionados a nuvem, na interrogação é aonde tem alguns tópicos de ajuda e também onde é possível entrar em contato com o suporte ou abrir um chamado para aumentar os limites de serviço da nuvem, no mundo é aonde o usuário consegue mudar o idioma da console da nuvem e por fim no símbolo de usuário o mesmo pode encontrar as informações dele.



Lab 2. Criando uma instância de Autonomous Data Warehouse





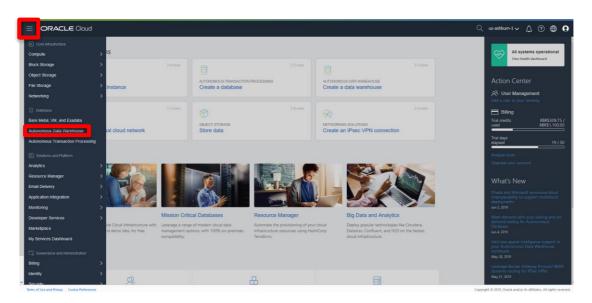
Lab 2. Criando uma instância de Autonomous Data Warehouse

Objetivos

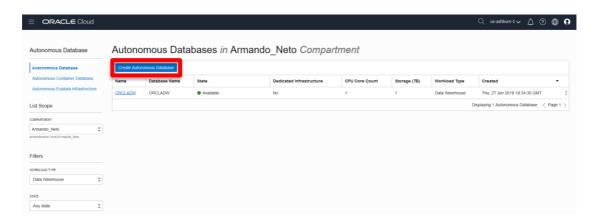
- Provisionar uma instância de Autonomous Data Warehouse de forma rápida
- Compreensão geral do Autonomous

Iniciando a criação da Instância de Autonomous Data Warehouse

Após feito o login no ambiente, serão acessados o menu e a opção Autonomous Data Warehouse, conforme a imagem abaixo.



Feito o passo acima a tela abaixo aparecerá.

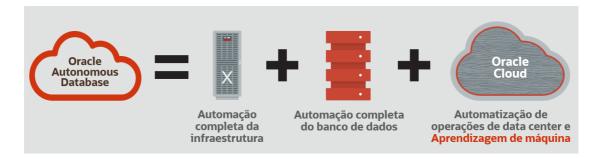


Feito o acesso, a primeira coisa que será feita é selecionar um Compartimento, no caso da imagem acima será acessado o compartimento "Armando_Neto", não é recomendado criar nenhum recurso no compartimento root, para a criação de um compartimento basta ir no menu, Identity e Compartments e criar um compartimento, mais detalhes neste link:

https://docs.cloud.oracle.com/iaas/Content/Identity/Tasks/managingcompartments.htm

Feito isso, como pode ser observado na imagem, é possível filtrar os bancos de dados autônomo por seu tipo de carga de trabalho e qual seu estado, ativo ou inativo por exemplo.

Além da possibilidade de criar uma instância de Autonomous *serverless*, que é a maneira que vai ser trabalhada neste laboratório, há a possibilidade de reservar um Autonomous Exadata Infrastructure e feito isso criar o Autonomous Container Database (CDB de Autonomous, é uma *feature* do banco Oracle chamada *Multitenant* introduzida na versão 12c), reservando o hardware e instanciando o CDB do Autonomous o usuário vai criar seu próprio banco Autonomous, feito isso, o usuário começa a criar os PDBs (Pluggable Databases) e como pode ser observado o Autonomous é um banco Oracle totalmente "stackado" como ilustra a imagem abaixo.

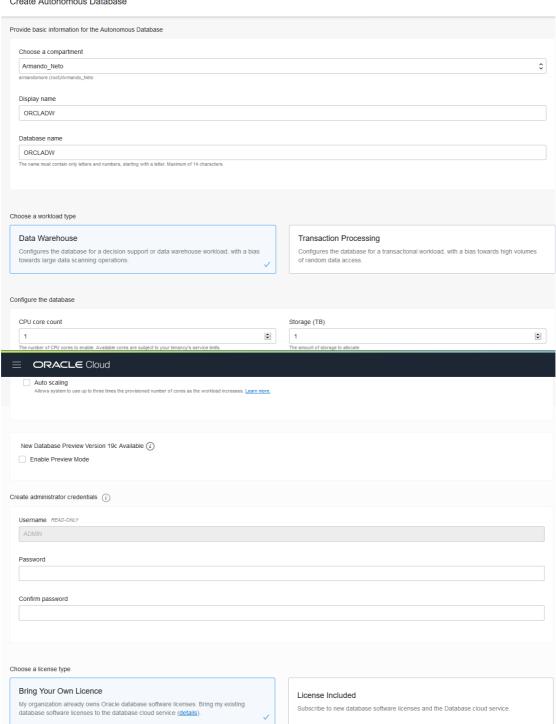


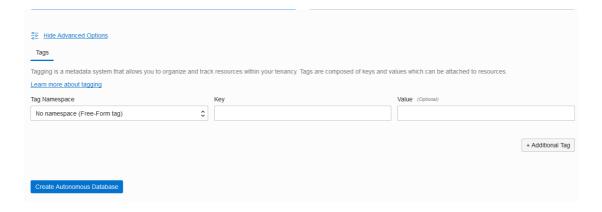
Como já mencionado, neste laboratório vamos trabalhar com o conceito serverless do Autonomous, clique no botão marcado na imagem "Create Autonomous Database".

Aparecerá uma lista de informações que devem ser preenchidas para a criação deste banco conforme as imagens abaixo.

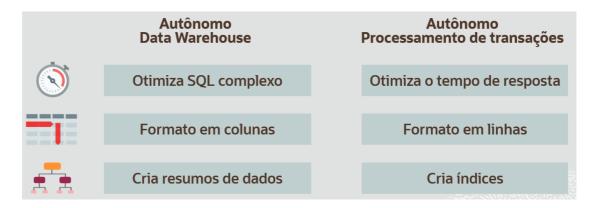
■ ORACLE Cloud

Create Autonomous Database





- Choose a compartment: Escolha o compartimento que deseja criar a instância do Autonomous, não é recomendado criar no root;
- Display Name: Aparecerá um nome gerado pela própria nuvem, mas pode ser colocado qualquer nome, será o nome dessa instância para a plataforma da nuvem. No caso da imagem foi colocado "ORCLADW";
- Database name: Nesta opção será o nome do banco (SID do banco), só pode conter letras e números e tem que começar o nome com uma letra. No caso da imagem foi colocado "ORCLADW";
- Choose a workload type: Aqui há duas opções, a Data Warehouse e a Transaction Processing, aqui o usuário escolhe qual Autonomous ele deseja, ele foca o banco para aprender e melhorar a performance ou de transações ou de queries analíticas, no caso deste laboratório será o workload de Data Warehouse. A imagem abaixo explica melhor as diferenças entre os workloads;



- Configure the database: Nesta parte será definido qual o poder de processamento e a quantidade de armazenamento será alocado, ambos são escaláveis tanto em processamento quanto em armazenamento, sem downtime para escalar, logo é recomendado começar com o mínimo que seria 1 OCPU de processamento e 1 Terabyte de armazenamento e ir escalando conforme vai crescendo as bases e o processamento;
- Auto scaling: Permitir o Auto scaling permite que o Autonomous multiplique por 3 o número de OCPU's que está inicialmente alocado a

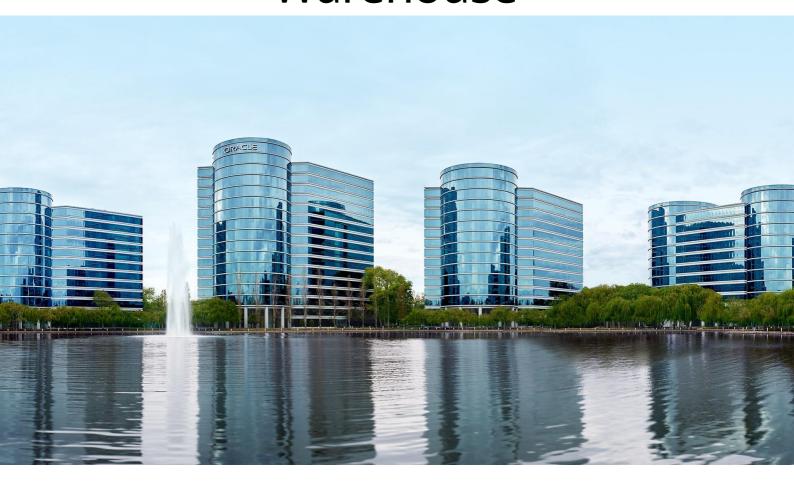
ele, por exemplo se for 1, ele irá escalar até 3, se for 2 ele vai escalar até 6 OCPU's, e assim sucessivamente. Ele irá escalar conforme ele percebe um aumento de carga, e ele escala gradualmente até atingir o seu máximo que é 3 vezes o inicialmente alocado, procurando sempre a melhor eficiência do processamento. Pode ser ativado e desativado a qualquer hora. O número de OCPU's que ele escalou será cobrado a mais pelo tanto de horas que ele ficou com mais OCPU's do que o normalmente alocado;

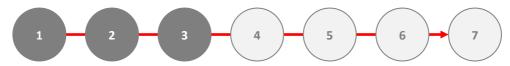
- New Database Preview Version 19c Available: Selecionando esta opção, ao invés do banco Autonomous usar a versão 18c como base ele usará a 19c:
- Create administrator credentials: O usuário com mais privilégios no Autonomous é o "ADMIN" e não pode ser alterado este Username, podendo criar mais usuários posteriormente quando a instância estiver pronta como qualquer outro banco, além disso a senha colocada aqui será a utilizada mais para a frente para acessar o banco, ela deve conter de 12 a 30 caracteres, uma maiúscula, uma minúscula e um número, não pode conter aspas ou escrito "admin";
- Choose a license type: Aqui há duas opções de licença, são elas:
 - License Included: É a licença da nuvem que será emprestada ao usuário já com suporte incluso e o mesmo paga preço de lista padrão;
 - Bring Your Own License (BYOL): Nesta modalidade o usuário traz a licença de banco Oracle local dele para a nuvem para pagar um preço diferenciado, para isto estar sempre em vigor o suporte da licença deve sempre estar ativo, mais detalhes podem ser encontrados neste link: https://cloud.oracle.com/en_US/datawarehouse/pricing
- Tags: Utilizado para monitoramento de custos de recursos ou projetos.

nos chats de vendas da Oracle:

Após preenchido tudo clique em "Create Autonomous Database", este processo de criação da Instância serverless do banco Autonomous demora por volta de 5 minutos.

Lab 3. Explorando o Autonomous Data Warehouse





Lab 3. Explorando o Autonomous Data Warehouse

Objetivos

- Overview do Autonomous Data Warehouse
- Conhecendo a console dos detalhes do Autonomous Data Warehouse
- Service Console

Overview do Autonomous Data Warehouse

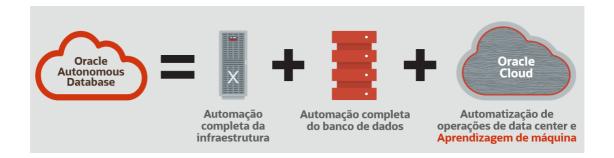
O Oracle Autonomous Data Warehouse fornece um banco de dados fácil de usar e totalmente autônomo que é escalável, oferece rápido desempenho de consultas e não exige administração de banco de dados, tudo isso sem **downtime**.

O Autonomous ele é independente, um serviço de data warehouse em nuvem totalmente gerenciado. Configuração de rede, armazenamento, aplicação de patches e upgrade em bancos de dados, além de outros serviços, são feitos pela Oracle.

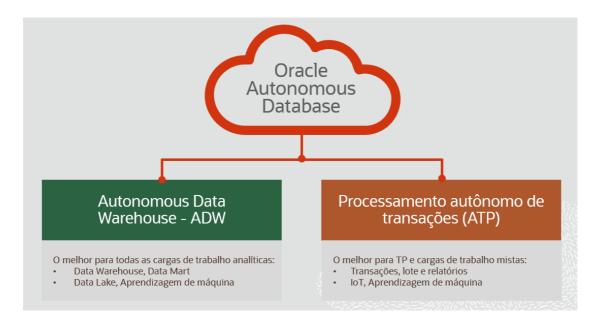
Ele se auto protege, sempre executa os patches de segurança mais recentes. Os dados armazenados são criptografados usando a criptografia TDE (Transparent Data Encryption). Os clientes de banco de dados usam SSL/TLS 1.2 criptografado e conexões mutuamente autenticadas.

Ele se autor repara, proteção automatizada contratempo de indisponibilidade. A alta disponibilidade está presente em cada componente, como por exemplo na hora de escalar ou de se aplicar um *patch* ele não terá *downtime*, e os backups são totalmente automatizados, podendo ser realizados backups manuais também.

Como já mencionado neste material, o Autonomous é um produto Oracle "stackado", ele está em uma infraestrutura Oracle que se chama Exadata que é desenhada para melhorar a performance de bancos Oracle, o sistema operacional deste hardware é o Oracle Linux com o Unbreakable Kernel que também é um sistema operacional otimizado para banco de dados e instalado neste sistema está o banco Oracle 18c, tudo isso aliado as automações e o Machine Learning da nuvem compõem o Autonomous Data Warehouse.

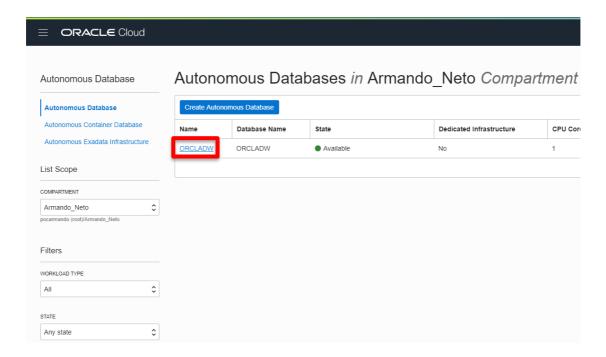


Além disso o Autonomous é uma família, o Data Warehouse é focado para processamentos analíticos e o Transaction Processing para transações e cargas de trabalho mistas, ótimo para desenvolvimento de aplicações que usam ele como base, segue abaixo uma imagem que representa essas soluções. Neste laboratório será abordado apenas o Autonomous Data Warehouse

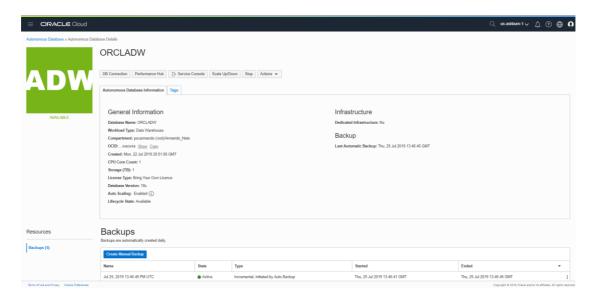


Acessando os detalhes do Autonomous Data Warehouse

Será acessado os detalhes do Autonomous que foi criado no passo anterior deste tutorial, primeiro clique no link conforme a imagem abaixo.



Feito o passo acima, será acessado a tela da imagem abaixo.



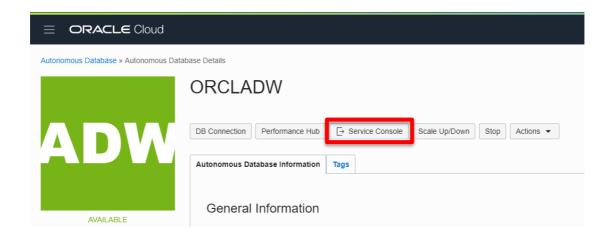
Nesta tela é possível ter as informações gerais do banco Autonomous e ver os backups que o próprio banco criou automaticamente, as regras automáticas dele é fazer backups diários incrementais com retenção de 60 dias, atualmente a retenção não pode ser alterada mas para mitigar isso o usuário pode criar os seus backups manuais e armazená-los em um Object Storage.

Pode-se observar que tem alguns botões no console também, e as ações deles são:

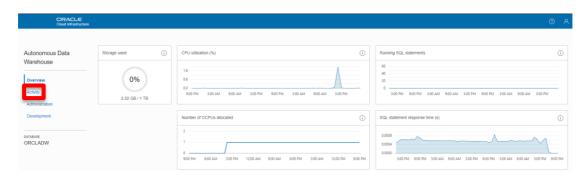
- DB Connection: Mostra uma janela com as Strings de conexão (High, Medium, Low) e permite que o usuário faça download das Credenciais do Cliente
- Performance Hub: Abre uma aba que monitora como está a performance no banco
- Service Console: Leva o usuário para a console de serviços, será explorado mais afundo no próximo tópico deste tutorial.
- Scale Up/Down: Abre uma janela que permite que o usuário escale o banco, tanto para cima quanto para baixo, tanto o armazenamento quanto o poder de processamento e todas essas alterações são independentes e **sem downtime**. Permite também o usuário ativar o Auto Scaling, *Feature* do banco que faz com que ele automaticamente aumente seu poder computacional em até 3 vezes, quando necessário, baseado no número de OCPU's que ele estava sem nenhum aumento automático.
- Stop: Para o serviço do banco e por consequência para o consumo de créditos relacionado ao poder computacional do banco, a cobrança do armazenamento continua.
- Actions
 - o Restore: Restaura um backup do banco
 - o Create Clone: Cria um clone do banco
 - Access Control List: Limita o acesso ao banco via um bloco de IPs (CIDR)
 - Admin Password: Altera a senha de ADMIN.
 - Update License Type: O usuário pode alterar o time de licença do banco sem precisar criar outro ou parar o mesmo.
 - Move Resource: Move o Autonomous para outro compartimento dentro do mesmo ambiente.
 - Apply Tag(s): Permite a inserção de tags no recurso para monitorar custos de projeto e afins.
 - Terminate: Deleta por completo o banco (A instância e o armazenamento).

Overview da *Service Console* do Autonomous Data Warehouse.

Após o overview dos detalhes do banco, será explorado a console de serviços do banco, para acessá-la, clique no botão conforme a imagem abaixo.

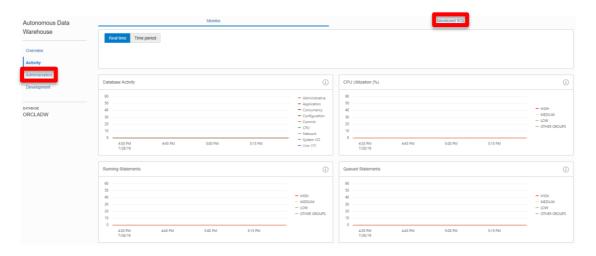


Feito o passo acima, o usuário irá acessar a seguinte console.



A console que aparece inicialmente logo que o usuário acessa a Service console é a de Overview, nela o usuário verifica como o Autonomous está sendo utilizado em poder computacional e armazenamento, quantas OCPU's alocadas e como está sua performance executando as queries.

Após a tela de overview será acessado a tela de "Activity", indicada na imagem acima.



Na tela de Activity o usuário monitora as atividades do Autonomous em tempo real, e pode chegar no nível de detalhe da query quando acessar a aba "Monitored SQL" indicada na imagem.

Após a familiarização, será acessado a tela "Administration" indicada na imagem acima.



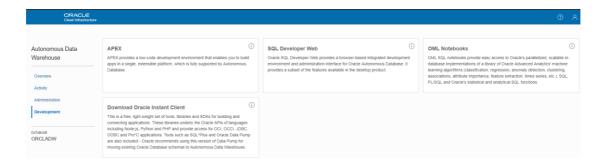
Nesta parte da console o usuário pode fazer o download das credenciais do Autonomous, da mesma maneira como mencionado anteriormente neste tutorial, o administrador do banco pode reconfigurar as regras dos recursos dedicados para cada tipo de conexão ou voltar as mesmas para o padrão, como já mencionado o Autonomous tem três maneiras de se conectar, elas são:

- "High": Conexão que permite paralelismo de queries, tem altos recursos dedicados para essa conexão e baixa concorrência.
- "Medium": Conexão que permite paralelismo de queries, menos recursos dedicados e mais concorrência entre os processamentos.
- "Low": Conexão que executa as queries em série, logo utiliza do paralelismo de queries existente no banco, poucos recursos alocados e alta concorrência entre os processos.

Ainda nesta console o usuário pode alterar a senha do login ADMIN do Autonomous, criar usuários do Oracle Machine Learning, que é uma ferramenta que permite os usuários executarem scripts de Machine Learning em SQL ou PL/SQL dentro do Autonomous utilizando do Machine Learning do mesmo, está em desenvolvimento a execução de scripts em R e Python, acessado via interface web é baseado em Apache Zeppelin.

E por fim, o usuário pode enviar um feedback para a Oracle, via aquele link, ele será redirecionado para um fórum onde ele não só pode enviar feedbacks como também compartilhar suas dúvidas e casos de sucesso.

Agora será acessado a tela de Development.

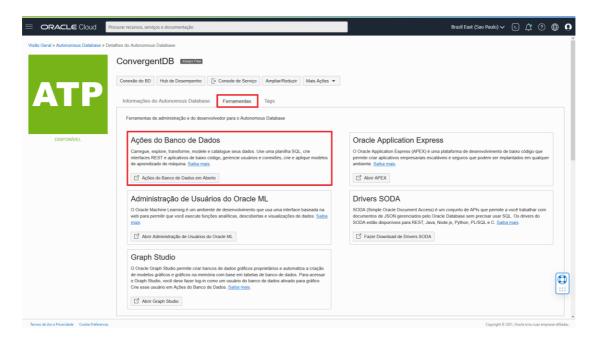


Nesta parte da console o usuário consegue acessar ferramentas de desenvolvimento no Autonomous, todas baseadas em Web, o usuário pode acessar:

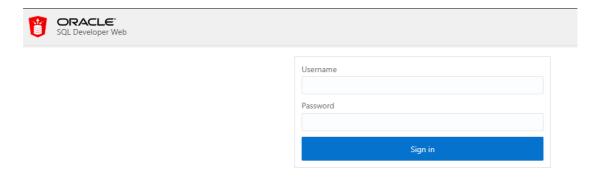
- APEX: O Oracle Application Express (APEX) é uma plataforma de desenvolvimento low-code que permite criar aplicativos incríveis, escaláveis em uma única plataforma que podem ser implementados em qualquer lugar, que é totalmente suportado pelo Banco de Dados Autonomous.
- SQL Developer Web: Provê um ambiente integrado de desenvolvimento e administração baseado em navegador web para o Oracle Autonomous. Ele provê uma parte das ferramentas disponíveis no produto desktop.
- Oracle Machine Learning: Os notebooks Oracle Machine Learning SQL fornecem acesso fácil às implementações escalonáveis e paralelizadas da Oracle em bancos de dados de uma biblioteca de algoritmos de aprendizado de máquina do Oracle Advanced Analytics (classificação, regressão, detecção de anomalias, agrupamentos, associações, importância de atributos, extração de recursos, séries temporais etc. .), SQL, PL / SQL e funções SQL analíticas e estatísticas da Oracle.
- Download Oracle Instant Client: O Oracle Instant Client permite que os aplicativos se conectem a um Banco de Dados Oracle local ou remoto para desenvolvimento e implementação de produção. As bibliotecas do Instant Client fornecem a conectividade de rede necessária, bem como recursos de dados básicos e de alto nível, para aproveitar ao máximo o Oracle Database. Ele é a base das APIs Oracle de linguagens e ambientes populares, incluindo Node.js, Python e PHP, além de fornecer acesso para aplicativos OCI, OCCI, JDBC, ODBC e Pro * C. As ferramentas incluídas no Instant Client, como o SQL * Plus e o Oracle Data Pump, fornecem acesso rápido e conveniente aos dados.

Executando queries no Autonomous Data Warehouse

O usuário pode executar queries no Autonomous via o SQL Developer Web. Para acessar este ambiente, na tela de detalhes do seu Autonomous, clique na aba "Ferramentas" e em seguida no botão "Ações do Banco de Dados".

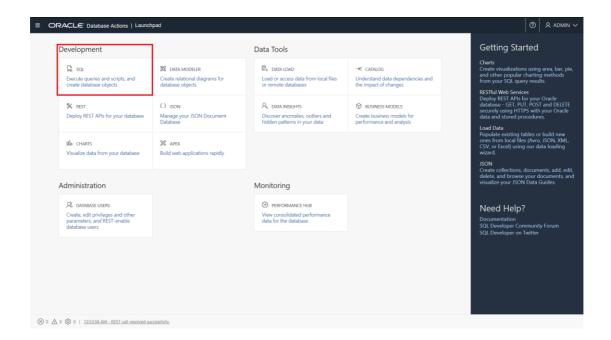


Inicialmente será acessado o SQL Developer Web, e aparecerá a tela de login.

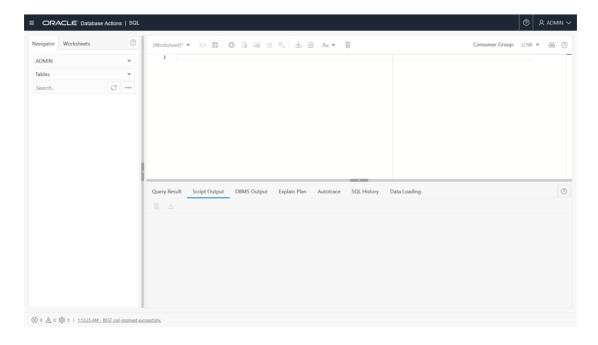


O login pode ser feito com o usuário ADMIN ou qualquer outro usuário que tenha sido criado no Autonomous. Neste caso o usuário utilizado será o ADMIN e a senha é a que foi utilizada na criação do Autonomous.

Feito isso será acessado a tela a "Ações do Banco de Dados". No menu de navegação, clique na opção "Development > SQL".



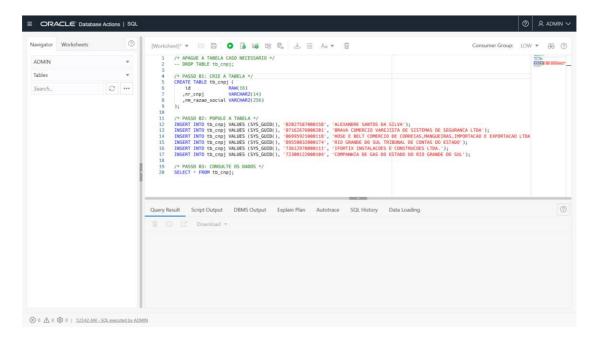
A seguinte tela do SQL Developer Web será acessada e poderemos executar as instruções SQL no banco de dados.



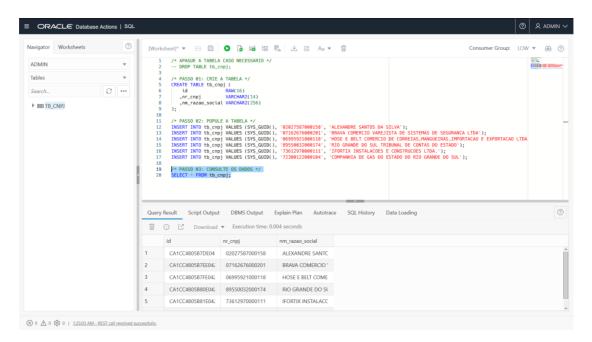
Carregando dados no Autonomous Data Warehouse

Será realizada uma carga de dados via SQL Developer Web para que seja trabalhado os Scripts de Banco de Dados convergente desejados.

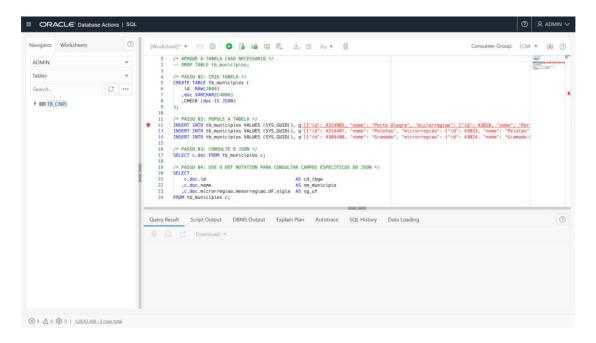
Inicie criando e inserindo dados na tabela de dados relacionais. Execute os scripts do arquivo "O1_PREPARACAO_TABULAR.sql", um passo por vez.



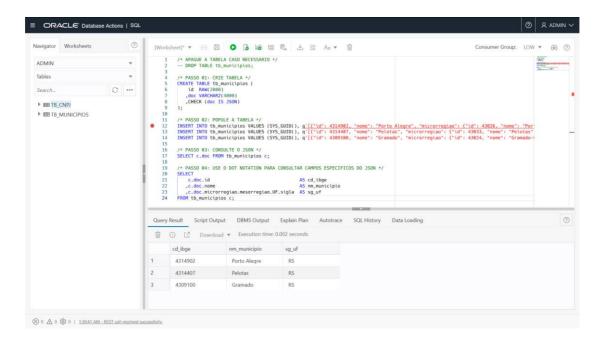
Ao executar o último passo, será possível consultar a tabela com os dados.



Em seguida, execute os scripts do arquivo "02_PREPARACAO_JSON.sql", um passo por vez. Não se preocupe com o aviso de "Unexpected Symbol" do editor, por se tratar de uma inserção de JSON, foi realizado um tratamento para inserir caracteres como aspas simples e aspas duplas.

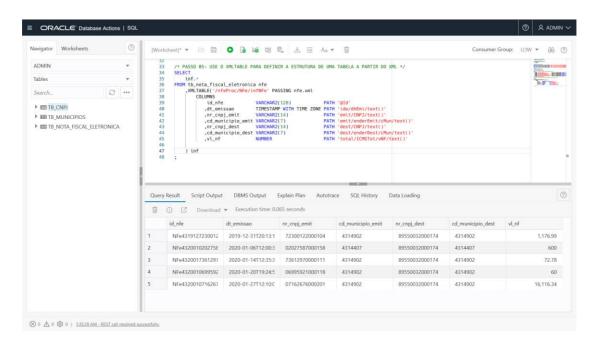


Ao executar o último passo, será possível consultar os dados dentro do documento JSON.

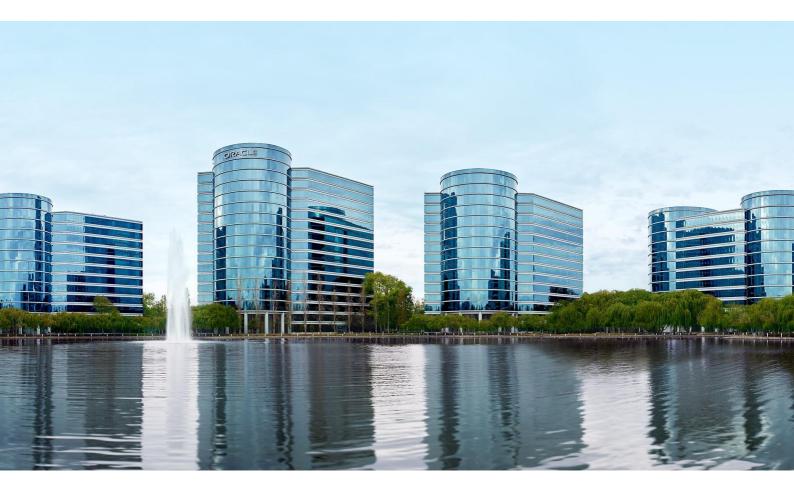


Por fim, execute os scripts do arquivo "03_PREPARACAO_XML.sql", um passo por vez.

Ao executar o último passo, será possível consultar os dados dentro do documento XML.



Lab 4. Consultando dados semi-estruturados no banco convergente





Lab 4. Consultando dados semiestruturados no banco convergente

Objetivos

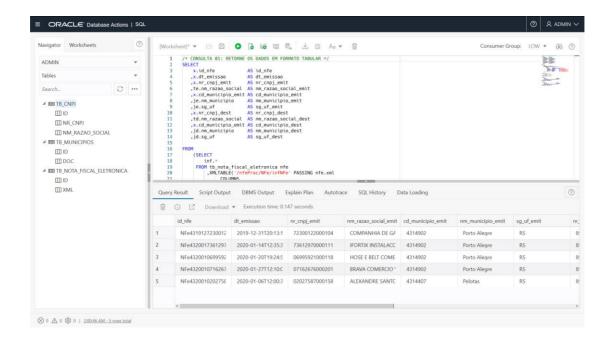
- Consultar os dados e retornar um resultado tabular
- Consultar os dados e retornar um resultado JSON
- Consultar os dados e retornar um resultado XML

Retornando dados em formato tabular

Ainda na tela do "SQL Developer Web", execute o script "04_CONSULTA_CONVERGENTE_TABULAR.sql". Note que dessa vez temos apenas um passo que consolida as consultas que realizamos nos scripts anteriores.

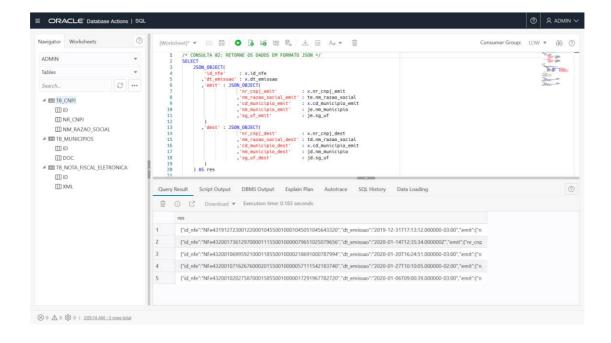
Na cláusula FROM do SQL, estamos acessando e relacionando as tabelas com dados relacionais, dados dos documentos JSON e dados dos documentos XML.

O resultado final nos é retornado em formato tabular.



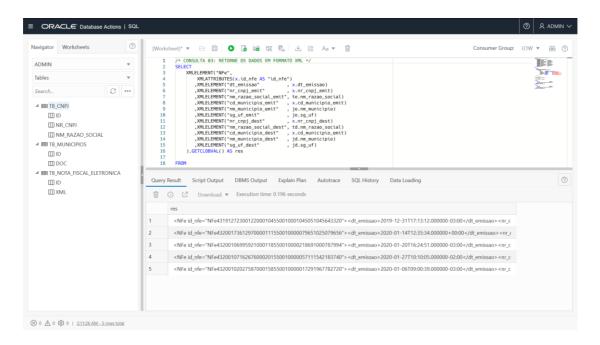
Retornando dados em formato JSON

Para visualizar um retorno em formato JSON, execute o script "05_CONSULTA_CONVERGENTE_JSON.sql". Note que dessa vez estamos modificando a cláusula SELECT e incluindo a função JSON_OBJECT para recebermos um retorno em formato JSON.



Retornando dados em formato XML

Por fim, para visualizar um retorno em formato XML, execute o script "06_CONSULTA_CONVERGENTE_XML.sql". Mais uma vez, fazemos ajustes na cláusula SELECT e utilizamos a função XMLELEMENT para recebermos um retorno em formato XML.



Parabéns! Você completou o Laboratório *Hands-On*

