



Inovação com dados em nuvem

**Acelerando com dados o
seu Dotinha**

Armando Moré Neto
Agosto 2021



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

Sumário

Pré-requisitos e Considerações Iniciais.....	4
Recursos usados:.....	4
Tópicos não cobertos.....	4
Considerações Iniciais.....	4
Objetivo e Ferramentas	6
Um pouco mais sobre o Dota 2.....	7
Acessando o Ambiente	12
Provisionando os Recursos	15
Banco de Dados Autonomous	15
Oracle Analytics Cloud	18
Carregando os dados no ADW.....	21
Conectando o OAC ao ADW.....	24
Preparação para a Análise Exploratória	28
Preparação dos conjuntos de dados	28
Executando a Análise Exploratória.....	40
Visão Geral	40
Performance dos Jogadores	48
Vitórias por Herói	50
Análise do Chat.....	52

Pré-requisitos e Considerações Iniciais

Recursos usados:

OCI

- Banco de dados Autonomous (ADW)
- Oracle Analytics Cloud (Consumo créditos do período de teste)

Local

- Navegador de sua escolha, no tutorial será usado o Chrome
- Oracle Analytics Desktop (Opcional, gratuito para desenvolvimento)

Kaggle

- Base de dados: <https://www.kaggle.com/devinanzelmo/dota-2-matches>

Tópicos não cobertos

- Como usar o Kaggle

Considerações Iniciais

Os videogames são uma área rica para extração de dados devido à sua natureza digital. Exemplos notáveis como a complexa economia de EVE Online, World of Warcraft e até mesmo os carros autônomos de Grand Theft Auto, nos diz que a ficção está mais perto da realidade do que realmente pensamos. Os cientistas de dados podem obter insights sobre a lógica e a tomada de decisões que os jogadores enfrentam quando colocados em cenários hipotéticos e virtuais.

O Conjunto de Dados utilizado: <https://www.kaggle.com/devinanzelmo/dota-2-matches>. Este conjunto de dados contém 50000 partidas ranqueadas do dump de dados do Dota 2 criado por Opendota.

Dota 2 é um MOBA popular disponível gratuitamente para jogar e pode levar milhares de horas da sua vida. O número de jogos neste conjunto de dados é jogado a cada hora. Se você gostar dos dados, há 2-3 milhões de correspondências adicionais facilmente disponíveis para download.

O objetivo deste conjunto de dados é permitir a exploração do comportamento do jogador, estimativa de habilidade ou qualquer coisa que você ache interessante. A

intenção é criar um recurso acessível e fácil de usar, que pode ser expandido e modificado se necessário.

O que está disponível na Base de Dados

- Matches: contém informações de alto nível sobre cada partida
- Players: jogadores individuais são identificados por account_id, mas há uma opção de jogar anonimamente e cerca de um terço do account_id não está disponível. Usuários anônimos têm o valor 0 para account_id. Contém totais para mortes, mortes, denies, etc. Contagens de ação do jogador estão disponíveis e são indicadas por nomes de variáveis começando com unit_order_. As contas de motivos para adquirir ou perder ouro e ganhar experiência têm os prefixos gold_ e xp_.
- player_time: contém os Last Hits, experiência e ouro amostrados em um intervalo de um minuto para todos os jogadores em todas as partidas. Os nomes das colunas indicam o player_slot. Por exemplo, xp_t_1 indica que esta coluna tem somas de experiência para o jogador no slot um.
- teamfights: tempo de início e fim de teamfights, bem como o tempo da última morte. As lutas de equipe parecem ser todas as batalhas com três ou mais mortes. Como tal, isso não inclui todas as batalhas da partida inteira.
- teamfights_players: informações adicionais fornecidas para cada jogador em cada teamfight. player_slot pode ser usado para vincular isso de volta a players.csv
- objetivos: Dá informações sobre todos os objetivos cumpridos, por que jogador e a que horas.
- chat: todos os bate-papos das 50 mil partidas. Há muitos palavrões e trollagens bem-humoradas.

Objetivo e Ferramentas

Neste tutorial, vamos trabalhar na criação de uma instância de Oracle Analytics Cloud e uma instância de Autonomous Data Warehouse seguindo processos e boas práticas e depois iremos explorar os dados de partidas ranqueadas do jogo Dota2 para tirarmos insights do jogo e aprender com os dados.

Explicando alguns dos recursos disponíveis no Oracle Analytics Cloud e no Autonomous Data Warehouse que são Plataformas como Serviço (PaaS). Ferramentas contempladas pelo Oracle Analytics Cloud:

- Analytics (Antigo Data Visualization, Versão Professional): Criação de visualizações interativas, ferramenta para transformações de bases de dados, análises preditivas, insights e forecasts, acesso aos aplicativos móveis Oracle Day by Day e Synopsis, para saber mais vide documentação.
 - Day by Day: <https://docs.oracle.com/en/cloud/paas/analytics-cloud/biday/index.html>
 - Synopsis: <https://docs.oracle.com/en/cloud/paas/analytics-cloud/biemy/index.html>
 - Oracle Analytics: <https://docs.oracle.com/en/cloud/paas/analytics-cloud/aaios/index.html>
- Classic [Antigo Oracle Business Intelligence Enterprise Edition (OBIEE), Versão Enterprise]: Criação de dashboards interativos, capacidade de gerar relatórios automáticos que podem ser enviados por e-mail, modelagem e gerenciamento de acessos as informações via Data Modeler.

Ferramentas fornecidas em conjunto com o Autonomous Data Warehouse:

- Oracle Analytics Desktop: Versão desktop do Analytics (Antigo Data Visualization) mencionado neste documento. Fornecido gratuitamente para o usuário que tem uma instância de Autonomous em sua Nuvem e para desenvolvimento.
- Database Actions: Ferramentas Web para gerenciamento dos Bancos de Dados Autônomo.
- Oracle Application Express: O Oracle Application Express (APEX) é uma plataforma de desenvolvimento de baixo código que permite criar aplicativos corporativos escaláveis e seguros, com recursos de classe mundial, que podem ser implantados em qualquer lugar.
- Oracle Machine Learning: O Oracle Machine Learning é uma interface de notebooks SQL e Python para que os cientistas de dados realizem o aprendizado de máquina no ADW (Oracle Autonomous Data Warehouse). Baseado em Apache Zeppelin.
- Graph Studio: O Graph Studio automatiza o gerenciamento de dados de grafos e simplifica a modelagem, análise e visualização em todo o ciclo de vida da análise de grafos.

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

Um pouco mais sobre o Dota 2

Dota 2 é um jogo eletrônico gratuito do gênero batalha multijogador, desenvolvido pela produtora Valve Corporation como sequência do Defense of the Ancients (DotA, uma modificação em um mapa desenvolvido para o jogo Warcraft III: The Frozen Throne), lançado em julho de 2013 na plataforma Steam.

Embora seu antecessor tenha sido desenvolvido de forma amadora, a versão 2 foi acolhido pela Valve, com parceria com o designer de jogos e produtor original IceFrog.

O jogo é baseado em partidas no modo online e offline, em cada uma delas o objetivo é derrotar a equipe adversária e, para isso, é, indispensavelmente, necessário destruir o Ancient, que, por sua vez, se localiza no centro da base inimiga.

Cada base é defendida por torres e ondas de “Creeps” que percorrem os caminhos principais do mapa (chamadas de “Lane”) que levam à suas bases. Existem 3 Lanes no mapa do jogo, que por convenção são chamadas de Top Lane (atravessa a parte de cima do mapa), Mid Lane (caminho que atravessa o centro do mapa) e Bot Lane (caminho que atravessa a parte de baixo do mapa).

No centro-norte do rio que corta o mapa, encontra-se o covil do Roshan, uma criatura que fica mais forte com o decorrer do jogo e que, ao ser derrotada, oferece recompensas, uma delas é o Aegis of the Immortal (um item que permite ao jogador renascer após sua morte) e o Queijo (item que permite ao jogador, depois de usá-lo em uma ação de "click" e recuperar 2500 pontos de vida e 1500 de mana) ou o Fragmento Restaurador (item que permite ao usuário reiniciar todos os seus contadores temporais de habilidades).

Ao centro de cada base está o Ancient, uma construção que deve ser destruída pela equipe adversária para que se ganhe o jogo, ele é dotado de autorregeneração. O mapa é dividido em duas áreas, sendo que uma delas pertence aos Iluminados (Radiant) e a outra aos Temidos (Dire). Os Radiant ficam localizados no canto sudoeste (canto inferior esquerdo) do mapa, enquanto os Dire ficam no canto nordeste (canto superior direito). As duas áreas do mapa são divididas por um rio que corre de uma ponta a outra cruzando a Mid Lane.

Há ainda áreas de floresta, "Jungle", onde existem criaturas neutras que podem ser subdivididas em comuns (com poucos pontos de vida), ou ancestrais (com muitos pontos de vida), sendo dois campos destas em cada lado do rio. Essas criaturas

garantem bônus de experiência e Gold ao jogador que as matar ou colaborar para isso (fazendo um acúmulo de criaturas, "stack", por exemplo).

Cada equipe é formada por exatamente 5 heróis, ou seja, totalizando 10 heróis participantes da partida. Caso um jogador abandone a partida, ele pode ser penalizado e jogar 5 partidas (ou mais) na "low priority", ou seja, partida de baixa prioridade para os servidores. Os jogadores que permanecem na partida em que um, ou mais jogadores abandonaram têm seus recursos aumentados, podem vender os itens do jogador que abandonou, além de poderem controlar o herói abandonado. O Dota 2 possui uma política de tratamento para os "maus jogadores", estes que podem, inclusive, ser reportados (por diferentes razões preestabelecidas) pelos participantes da partida.

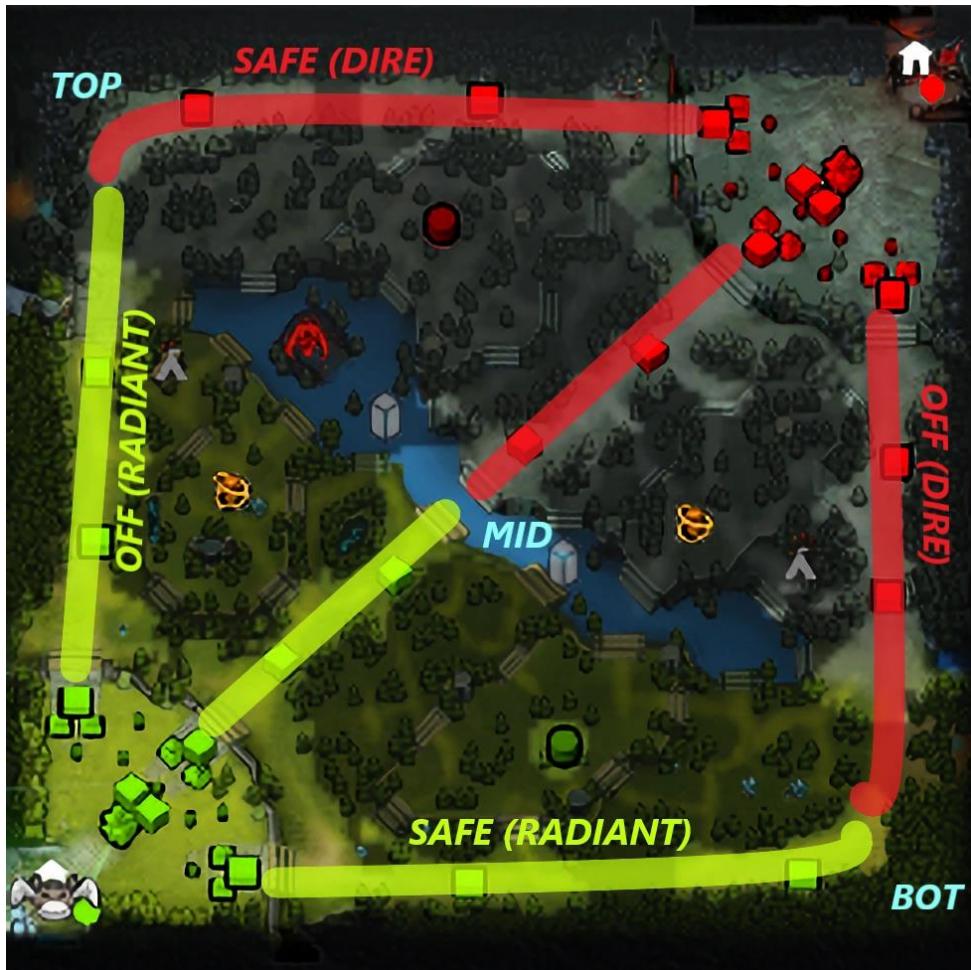
Cada jogador controla um personagem, chamado de "herói" (há, atualmente, 115 heróis disponíveis, com o possível lançamento de mais), com o qual participa dos combates que, geralmente, acontecem ao longo das três Lanes, ou na Jungle.

Baseado em elementos de RPG Eletrônico de Ação, os heróis iniciam no nível 1, podendo chegar até o nível 25. Os heróis avançam de nível conforme ganham pontos de experiência. A cada nível avançado, os heróis recebem 1 ponto para ser usado para melhorar suas habilidades especiais ou acrescentar talentos, em níveis específicos. Então, quanto maior o nível do herói, melhor são seus talentos e maior será a potência ou utilidade de suas habilidades. Os elementos do jogo são amplamente incorporados do seu antecessor original (Dota), com seu desenvolvedor, IceFrog, sendo o atual e mais antigo desenvolvedor deste.

Vale lembrar, ainda, que há vários modos de partidas, tipos diferentes de jogo, dentro do Dota 2, seja os tipos de partida competitiva (ranqueada) ou livre, seja como "Seleção de Habilidades", "Turbo", entre outros.

Em maio de 2013, o Dota 2 chegou a quase 330 mil jogadores simultâneos e detém o recorde para o jogo com mais usuários simultâneos na história da Steam, quebrando seu próprio recorde estabelecido em março do mesmo ano. Simultaneamente com esta referência, que determinou-se que o número de jogadores simultâneos em maio 2013 superou o número de jogadores para o resto dos dez jogos mais jogados da Steam combinados.

Lanes



As Lanes são uma das principais características do mapa. Existem três Lanes distintas, cada uma com suas próprias vantagens. Eles são identificados por meio de várias abordagens táticas diferentes.

- Top Lane: Geralmente abreviado para Top é a Lane que vai de base a base por meio do canto superior esquerdo do mapa. Ela corre ao longo dos lados esquerdo e superior do mapa. Tanto o Radiant quanto o Dire se referem a esta via como a via superior. É a Radiant Offlane e a Dire Safe Lane. É adjacente à selva dos Dire e a onda de creep frequentemente acaba lutando ao lado dos Dire.
- Bot Lane: Geralmente abreviado para Bot, é a Lane que vai de base a base por meio do canto inferior direito do mapa. Ele corre ao longo dos lados inferior e direito do mapa. Tanto o Radiant quanto o Dire se referem a esta via como bot lane. É a Dire Offlane e a Radiant Safe Lane. É adjacente à selva do Radiant e a onda de creep frequentemente acaba lutando ao lado do Radiant.

- Mid Lane: Geralmente abreviada para Mid é a rota que vai diretamente de uma base, através do rio, até a base do oponente. É o caminho mais rápido e mais curto para cada base. O Mid é tradicionalmente uma via solo. Envolve muita habilidade, geralmente envolvendo o melhor jogador de uma equipe. O mid laner deve contestar os últimos Last Hits e Denies para controlar o fluxo de ouro e exp a seu favor. Isso tem a vantagem de controlar o poder do mid laner inimigo, enquanto também reforça sua própria força. O jogador que "vence" a rota intermediária é essencialmente capaz de gankar as rotas laterais com sucesso, bem como empurrar para baixo a torre da rota intermediária do inimigo, ganhando o controle do mapa para todo o seu time. Muito do jogo na Mid Lane também gira em torno do controle de runas e evitando ganks.
- Safe Lane: A Safe Lane é a Lane adjacente à sua Jungle. Por ser mais longa, seus creeps levarão mais tempo para se reforçar, mantendo a onda creep mais perto de sua torre e do seu lado do rio. Como resultado, os heróis estarão mais seguros contra ganks, pois podem ficar sob sua torre. Se sua equipe tem um jungler, ele também estará por perto em caso de emergências e uma posição perto da torre. Para o lado Radiant, a pista inferior é a pista segura, rodando ao longo da parte inferior do mapa. Para o lado Dire, a pista superior é a pista segura, correndo ao longo do topo do mapa.
- Offlane: A Offlane é adjacente à sua loja secreta e aos antigos acampamentos. Devido à curta distância de reforço, seus creeps muitas vezes acabam no lado do rio do oponente, deixando você em uma posição vulnerável próximo à selva dele. Ser suscetível a heróis vindo pelo rio torna esta rota mais desafiadora, entretanto, estrategicamente, o Offlaner é normalmente executado sozinho para que o Offlaner ganhe EXP rapidamente para deixar a rota no nível 6 e gankar. O nome "Offlane" refere-se ao estilo de jogo de correr na lane, no qual o herói raramente é visto pela última vez a acertar na lane. Em vez disso, eles estão "Offlane" na distância mais segura possível para ganhar experiência.

Algumas estratégias usadas nas Lanes do Dota2:

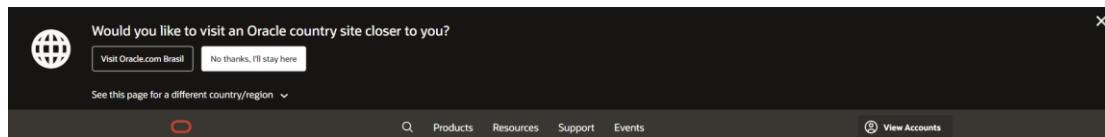
- Solo Lane: É quando uma pista tem apenas um único herói correndo. Os heróis que estão sozinhos em uma rota recebem o máximo de experiência e ouro, pois não é dividido entre os ocupantes. O principal risco dessa estratégia recompensadora é que ficar sozinho torna o herói mais vulnerável a ganks. Na estratégia comum, a pista do meio é sempre executada sozinha e a Offlane geralmente é executada sozinha.

- Dual Lane: É quando dois heróis se unem para correr uma pista juntos. Esta é a configuração mais comum para a via segura. Normalmente, um herói que é mais dependente de farm (Ato de dar Last Hits nas Creeps) tenta maximizar seus últimos golpes, enquanto o segundo herói apóia o outro protegendo-o, agredindo o oponente e obtendo Denies (Ato de dar Last Hit nas Creeps do seu Time antes que o inimigo o faça).
- Tri Lane: É quando três heróis correm uma pista juntos. Existem dois tipos de estratégias de trilane.
 - O trilane agressivo é três heróis na Offlane. Eles tentam pressionar a rota segura do time inimigo e pegar a Jungle do oponente. Para ser bem-sucedido, um trilane agressivo depende muito de matar o herói ou a torre.
 - O trilane defensivo consiste em três heróis na Safe Lane. Dois heróis coadjuvantes garantem que o terceiro herói, seu carry, pegue todo o farm. Eles empurram os heróis inimigos de volta para negar-lhes experiência e ouro, ao mesmo tempo que protegem seu carry do inimigo. Os heróis coadjuvantes podem ganhar experiência e ouro próprios farmando a selva enquanto isso.

A palavra trilane às vezes também pode se referir à configuração de trilane, referindo-se a uma divisão 1-1-3 de heróis entre as raias. Nesse caso, a equipe conta com duas pistas solo e uma trilane.

Acessando o Ambiente

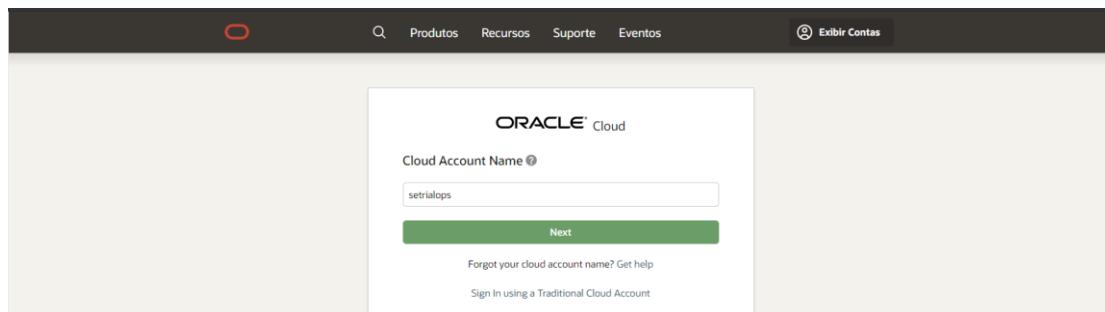
Vá para oracle.com. Você pode alterar o idioma dessa página antes do acesso ao ambiente:



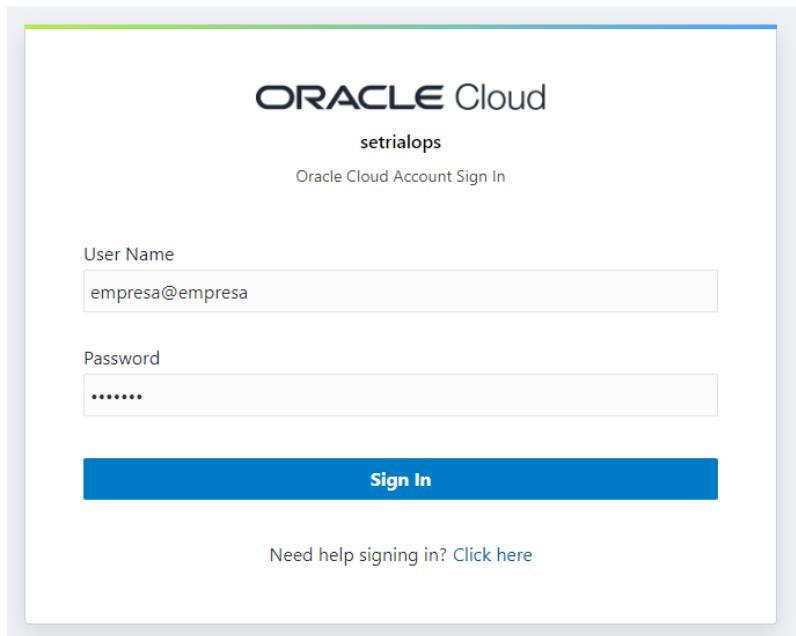
No site já em português, clique em Exibir Contas e depois em Faça Login na Nuvem:



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.

The screenshot displays the Oracle Cloud homepage with the following sections:

- Quick Actions:**
 - COMPUTE: Create a VM instance (0.4 mins)
 - AUTONOMOUS TRANSACTION PROCESSING: Create a database (0.4 mins)
 - AUTONOMOUS DATA WAREHOUSE: Create a data warehouse (0.4 mins)
 - NETWORKING: Create a virtual cloud network (0.4 mins)
 - OBJECT STORAGE: Store data (0.4 mins)
 - NETWORKING SOLUTIONS: Create an IPsec VPN connection (0.4 mins)
- Solutions:**
 - Jump Start:** Hands-on use of Oracle Cloud Infrastructure with self-paced learning and demo labs, for free.
 - Mission Critical Databases:** Leverage a range of modern cloud data management options, with 100% on-premises compatibility.
 - Resource Manager:** Automate the provisioning of your cloud infrastructure resources using HashiCorp Terraform.
 - Big Data and Analytics:** Deploy popular technologies like Cloudera, Datastax, Confluent, and Hadoop on the fastest cloud infrastructure.
- Learn:**

Adding users and groups	Understanding compartments	Launching your first instance
Understanding service limits	Key concepts and terminology	Overview of Autonomous Database
- Action Center:**
 - User Management:** Add a user to your tenancy.
 - Billing:** Total credits used: 4885215.28 / RRs1,100.00. Total days elapsed: 15 / 30. Accept costs. Upgrade your account.
 - What's New:**
 - Oracle and Microsft announced cloud partnership to support microservices deployments (May 2, 2019)
 - Mount volumes with public catalog and no-deleted training for Autonomous Database (May 4, 2019)
 - Add new spatial intelligence support to your Autonomous Data Warehouse (May 20, 2019)
 - Leverage Border Gateway Protocol (BGP) for load balancing for IPsec VPNs (May 21, 2019)
 - Minimize failure deployments by using Oracle Cloud Infrastructure StackSets for AutoMates (May 28, 2019)
 - Get Help:** Contact Support, Developer Tools, Documentation.

Small text at the bottom left: Terms of Use and Privacy | Cookie Preferences

Small text at the bottom right: Copyright © 2019, Oracle and/or its affiliates. All rights reserved.

Provisionando os Recursos

Banco de Dados Autonomous

Navegue no menu direito até Autonomous Data Warehouse e clique em criar Autonomous Database. Para esse tutorial será usado o ADW (Data Warehouse).

The screenshot shows the Oracle Cloud interface for creating an Autonomous Database. The top navigation bar includes the Oracle Cloud logo and a search bar. The main section is titled "Criar Autonomous Database" (Create Autonomous Database). It starts with a step to "Fornecer informações básicas para o Autonomous Database" (Provide basic information for the Autonomous Database). This step includes fields for "Compartimento" (Compartment), "Nome para exibição" (Display name), and "Nome do banco de dados" (Database name). Below this, there's a section to "Escolher um tipo de carga de trabalho" (Choose a type of workload). It lists three options: "Data Warehouse" (selected), "Processamento de Transações" (Transaction Processing), and "JSON". Each option has a brief description. Next, it asks to "Escolher um tipo de implantação" (Choose a deployment type), with "Infrastructure Compartilhado" (Shared Infrastructure) selected. A note indicates that Exadata Infrastructure Dedicated is not available for Oracle Autonomous Database Always Free. Then, it moves to "Configure o banco de dados" (Configure the database). It shows "Always Free" selected, with a note about automatic shutdown after 7 days. It also shows "19c" as the database version, "1" for CPU cores, and "0.02" for storage. Finally, there are "Criar Autonomous Database" and "Cancelar" (Cancel) buttons at the bottom.

Criar Autonomous Database

Escolher versão do banco de dados

19c

Contagem de OCPUs Somente para Leitura
1

Os Autonomous Databases Always Free podem utilizar até 1 núcleo. A contagem de núcleos de CPU não pode ser ajustada.

Escalonamento automático
 Permite que o sistema use até três vezes o número provisionado de núcleos à medida que a carga de trabalho aumenta. [Saiba mais](#)

Armazenamento (TB) Somente para Leitura
0.02

Os bancos de dados Autônomos Always Free podem utilizar até 0,02 TB (20 GB) de armazenamento. O tamanho do armazenamento não pode ser ajustado.

Crie credenciais de administrador

Nome do usuário Somente para Leitura
ADMIN

O nome de usuário ADMIN não pode ser editado.

Senha

Confirme a senha

Escolher acesso à rede

Tipo de Acesso

Permitir acesso seguro de todos os lugares
Você pode restringir o acesso a endereços IP e VCNs específicos.

Rede virtual na nuvem
Somente acesso privado, usando uma VCN.

Configurar regras de controle de acesso [\(i\)](#)

Escolha um tipo de licença

BYOL (Bring Your Own License)
Trazer as licenças de software do Oracle Database da minha organização para o serviço de Banco de dados. [Saiba mais](#)

Licença Incluída
Assine novas licenças de software do Oracle Database e o serviço de Banco de Dados.

[Mostrar Opções Avançadas](#)

Criar Autonomous Database **Cancelar**

[Termos de Uso e Privacidade](#) [Cookie Preferences](#)

EXEMPLO [Always Free](#)

Informações do Autonomous Database

ADW **PROVISIONANDO**

Informações Gerais

Nome do Banco de Dados: EXEMPLO
Tipo de Carga de Trabalho: Data Warehouse
Compartimento: **señihritis (root) /racz/Armando**
OCID: **575x9x** [Mostrar Código](#)
Criado: ter, 16 de mar de 2021 17:00:04 UTC
Contagem de OCPUs: 1
Armazenamento: 20 GB
Tipo de Licença: Licença incluída
Versão do Banco de Dados: 19c
Dimensionamento Automático: Desativado [\(i\)](#)
Estado do ciclo de vida: Provisionando
Tipo de Instância: Gratuíto [Upgrade para Pago](#)
Modo: Leitura/Gravação [Editar](#)

Infraestrutura

Infraestrutura Dedicada: Não
Autonomous Data Guard [\(i\)](#)
Status: Desativado [\(i\)](#)

Fazer Backup

Último Backup Automático: Não há backups ativos para este banco de dados
Armazenamento de Backup Manual: Não Configurado

Rede

Tipo de Acesso: Permitir acesso seguro de todos os lugares
Lista de Controle de Acesso: Desativado [Editar](#)

Métricas

Recursos

Métricas

Horário de início: 16 de mar de 2021 15:41:53 Horário de término: 16 de mar de 2021 16:41:53 Seleções Rápidas: Última hora

Utilização de CPU [\(i\)](#)
Intervalo: 1 minuto Estatística: Maior

Utilização do Espaço de Armazenamento [\(i\)](#)
Intervalo: 1 hora Estatística: Maior

Copyright © 2021, Oracle e/ou suas empresas afiliadas. Todos os direitos reservados.

The screenshot shows the Oracle Cloud interface for an Autonomous Database named 'EXEMPLO'. The top navigation bar includes 'Procurar recursos, serviços e documentação', 'Visão Geral', 'Autonomous Database', and 'Detalhes do Autonomous Database'. The main content area is titled 'EXEMPLO Always Free'.

Informações do Banco de Dados

- Nome do Banco de Dados: EXEMPLO
- Tipo de Carga de Trabalho: Data Warehouse
- Compartimento: serialldha (raiz) /Armando
- OCID: ...57sxkx [Mostrar](#) [Copiar](#)
- Criado: ter., 16 de mar. de 2021 17:00:04 UTC
- Contagens de OCPIUs: 1
- Armazenamento: 20 GB
- Tipo de Licença: Licença incluída
- Versão do Banco de Dados: 19c
- Dimensionamento Automático: Desativado [\(i\)](#)
- Estado do ciclo de vida: Disponível
- Tipo de Instância: Gratuito [Upgrade para Pago](#)
- Modo: Leitura/Gravação [Editar](#)

Infraestrutura

- Infraestrutura Dedicada: Não
- Autonomous Data Guard [\(i\)](#)
- Status: Desativado [\(i\)](#)
- Fazer Backup
- Último Backup Automático: Não há backups ativos para este banco de dados.
- Armazenamento de Backup Manual: Não Configurado

Rede

- Tipo de Acesso: Permitir acesso seguro de todos os lugares
- Lista de Controle de Acesso: Desativado [Editar](#)

Manutenção [\(i\)](#)

- Próxima Manutenção: sáb., 20 de mar. de 2021 4:00:00 UTC - 8:00:00 UTC

Recursos

- Métricas
- Backups (0)
- Clones Atualizáveis (0)
- Solicitações de Serviço (1)

Métricas

Horário de início: 16 de março de 2021 15:41:53 | Horário de término: 16 de março de 2021 16:41:53 | Seleções Rápidas: Última hora

Utilização de CPU [\(i\)](#)

Intervalo	1 minuto	Estatística	Média
100	80		

Opções [\(i\)](#)

Utilização do Espaço de Armazenamento [\(i\)](#)

Intervalo	1 hora	Estatística	Média
100	80		

[Termos de Uso e Privacidade](#) [Cookie Preferences](#)

Para saber mais sobre o banco de dados Autônomo, visite os links abaixo:

- Página do produto: <https://www.oracle.com/br/autonomous-database/>
- Documentação: <https://docs.oracle.com/en/cloud/paas/autonomous-data-warehouse-cloud/>
- Laboratório Introdutório (ADW):

https://apexapps.oracle.com/pls/apex/f?p=44785:50:105649591877528::50:P50 COURSE_ID,P50_EVENT_ID:251,5925

Oracle Analytics Cloud

Navegue no menu direito até Análise e Analytics Cloud clique em Criar Instância. Para este tutorial será usado a versão Self-Service Analytics (Conhecida também como versão Professional) com 2 OCPUs.

The screenshot shows the Oracle Cloud interface with the sidebar navigation open. The 'Analytics' section is selected, and 'Analytics Cloud' is highlighted. Under 'Analytics Cloud', the 'Flexo' option is selected. The main content area displays a table titled 'Instâncias de Análise em setrialItda (root) (raiz) Compartimento'. The table has columns for 'Nome', 'Status', 'Conjunto de funcionalidades', 'Capacidade', and 'Criado'. A message at the bottom of the table says 'Nenhum item encontrado.' and 'Mostrando 0 itens 1 de 1'.

ORACLE Cloud

Criar Instância de Análise

Informações Básicas

Compartimento
Armando
semtalida (root) (raiz)/Armando

Nome da instância

Descrição Opcional

Recursos

Conjunto de funcionalidades

Capacidade
OCPUs

Licenciamento (i)

Licença incluída	BYOL (Bring Your Own License)
Assinar uma nova licença do software Analytics Cloud e o serviço Analytics Cloud. Saiba mais. ✓	

Acesso à Rede

Público Acesso de qualquer lugar Privado Acesso apenas de uma Rede Virtual na Nuvem

Configurar Regras de Controle de Acesso

Tags

Tag é um sistema de metadados que permite organizar e rastrear recursos na sua tenancy. As tags são compostas de chaves e valores que podem ser anexados a recursos. [Saiba mais sobre tags](#)

Namespace de Tag	Chave de Tag	Valor

Ações [Criar](#) [Cancelar](#)

Depois de clicar em Criar conforme a imagem acima o Oracle Analytics Cloud começará a provisionar e este processo demora entre 15 a 30 minutos.

ORACLE Cloud Brazil East (Sao Paulo) A instância do Analytics EXEMPLANOANALYTICS foi criada com sucesso

Análise > Instâncias de Análise > Detalhes da Instância

EXEMPLANOANALYTICS

 CRIANDO

[Detalhes da Instância](#) [Detalhes Adicionais](#) [Tags](#)

Informações Básicas

OCID: [mjdjzde4w](#) Master_Costar
Compartimento: semtalida (root) (raiz)/Armando
Criado: ter, 16 de mar. de 2021 17:15:40 UTC
Capacidade: 2 OCPUs
Conjunto de funcionalidades: Self-Service Analytics
Tipo de Licença: BYOL (Bring Your Own License)

Informações de Acesso

URL: Nenhum Valor
URL Personalizada: Nenhum Valor

Log de Atividade

Ação	Status	Início	Duração
CREATE_ANALYTICS_INSTANCE	ACCEPTED		

Mostrando 1 item < 1 de 1 >

The screenshot shows the Oracle Cloud interface for managing analytics instances. At the top, there's a navigation bar with 'ORACLE Cloud' and a search bar. Below it, the path 'Análise > Instâncias de Análise > Detalhes da Instância' leads to the 'EXEMPLANOANALYTICS' instance. On the left, there's a green hexagonal icon with 'OAC' and the word 'ATIVO' below it. The main content area has tabs for 'Detalhes da Instância', 'Detalhes Adicionais', and 'Tags'. Under 'Informações Básicas', it shows the OCID, compartment ('setrialtda (racy)/Armando'), creation date ('ter., 16 de mar. de 2021 17:15:40 UTC'), and capacity ('2 OCPUs'). It also indicates 'Conjunto de funcionalidades: Self-Service Analytics' and 'Tipo de Licença: BYOL (Bring Your Own License)'. In the 'Rede' section, it says 'Tipo de Acesso: Público' and 'Regras de Controle de Acesso: Não Configurado'. The 'Informações de Acesso' section provides the URL ('https://exemplanoanalytics-5ewxboipgud.gr...') and a 'URL Personalizado' link. Below this, a 'Log de Atividade' table lists a single entry: 'CREATE_ANALYTICS_INSTANCE' with status 'SUCCEEDED', initiated on 'ter., 16 de mar. de 2021 17:15:51 UTC', and a duration of '28 min 19 s'. The bottom of the page includes links for 'Termos de Uso e Privacidade' and 'Cookie Preferences', and a copyright notice: 'Copyright © 2021, Oracle e/ou suas empresas afiliadas. Todos os direitos reservados.'

Para acessar o ambiente do Oracle Analytics Cloud basta clicar em Home Page do Analytics.

Para conhecer mais sobre o Oracle Analytics, visite os links abaixo:

Página do produto: <https://www.oracle.com/br/business-analytics/analytics-cloud.html>

Documentação: <https://docs.oracle.com/en/cloud/paas/analytics-cloud/index.html>

Carregando os dados no ADW

Agora que tanto o ADW quanto o OAC estão criados e disponíveis, serão carregados os conjuntos de dados que serão usados neste tutorial.

Serão usados os arquivos abaixo do link do Kaggle (<https://www.kaggle.com/devinanzelmo/dota-2-matches>) que já foi mencionado neste documento, são eles:

1. chat.csv
2. hero_names.csv
3. match.csv
4. players.csv

Dentro do ADW serão carregados os arquivos csv's, para isso acesse os detalhes do Autonomous, vá em ferramentas e clique em Ações de Bancos de Dados em Aberto e siga o passo a passo abaixo.

The screenshot shows the Oracle Cloud Autonomous Database management interface. At the top, there's a navigation bar with 'ORACLE Cloud' and a search bar. Below it, the main title is 'Visão Geral » Autonomous Database » Detalhes do Autonomous Database'. On the left, there's a large green box labeled 'ADW' and 'DISPONÍVEL'. The main content area has several sections:

- EXEMPLO Always Free**: Includes tabs for 'Conexão do BD', 'Hub de Desempenho', 'Console de Serviço', 'Ampliar/Reduzir', and 'Mais Ações'. The 'Mais Ações' tab is currently selected.
- Informações do Autonomous Database**, **Ferramentas**, and **Tags** buttons.
- Ferramentas de administração e do desenvolvedor para o Autonomous Database**:
 - Ações do Banco de Dados**: Describes how to manage data using SQL, REST APIs, and machine learning. It includes a button 'Ações do Banco de Dados em Aberto'.
 - Oracle Application Express**: Describes APEX as a low-code development platform. It includes a button 'Abrir APEX'.
 - Administração de Usuários do Oracle ML**: Describes Oracle Machine Learning. It includes a button 'Abrir Administração de Usuários do Oracle ML'.
 - Drivers SODA**: Describes Simple Oracle Document Access. It includes a button 'Fazer Download de Drivers SODA'.
- Recursos** and **Métricas** sections.
- Database Actions** section: Shows a login form with 'Username' (ADMIN) and 'Password' (redacted). A 'Next' button is visible.
- Database Actions** section: Shows a login form with 'Username' (ADMIN), 'Password' (redacted), and a 'Sign in' button.

The screenshot shows the Oracle Database Actions interface. In the top right corner, there are user icons and a dropdown menu labeled "ADMIN". The main area is divided into several sections:

- Development** section:
 - SQL**: Execute queries and scripts, and create database objects.
 - DATA MODELER**: Create relational diagrams for database objects.
 - REST**: Deploy REST APIs for your database.
 - JSON**: Manage your JSON Document Database.
 - APEX**: Build web applications rapidly.
- Data Tools** section:
 - DATA LOAD**: Load or access data from local files or remote databases.
 - CATALOGS**: Understand data dependencies and the impact of changes.
 - DATA INSIGHTS**: Discover anomalies, outliers and hidden patterns in your data.
 - BUSINESS MODELS**: Create business models for performance and analysis.
- Administration** section:
 - DATABASE USERS**: Create, edit, and REST enable database users.
- Getting Started** section:
 - RESTful Web Services**: Deploy REST APIs for your Oracle database - GET, PUT, POST and DELETE securely using HTTPS with your Oracle data and stored procedures.
 - Load Data**: Import existing tables or build new ones from local files (Avro, JSON, XML, CSV, or Excel) using our data loading wizard.
 - JSON**: Create collections, documents, add, edit, update, and browse your documents, and visualize your JSON Data Models.
 - Available On-Premises**: SQL Developer Web is now available for your On-Premises Oracle Databases tool!
- Need Help?** section:
 - Documentation
 - SQL Developer Community Forum
 - SQL Developer on Twitter

This screenshot shows the "Data Load" step of the Oracle Database Actions process. The top navigation bar includes "ORACLE Database Actions", "Data Load", and "ADMIN". The main content area is titled "What do you want to do with your data?" and contains three options:

- LOAD DATA**: Import data into your Autonomous Database.
- LINK DATA**: Leave your data in place and let your Autonomous Database access it.
- FEED DATA**: Setup ongoing feed of new data into your Autonomous Database.

Below this, the question "Where is your data?" is asked, with three options:

- LOCAL FILE**: Select text or Excel files from your local device.
- DATABASE**: Select tables from your remote databases.
- CLOUD STORAGE**: Select buckets from cloud storage (Oracle, S3, Azure, Google).

A "Next" button is located at the bottom right of this section. Further down, under "Explore and Connect", there are two more options:

- EXPLORE**: Inspect data in your Autonomous Database.
- CLOUD LOCATIONS**: Manage connections to your cloud storage (Oracle, S3, Azure, Google).

This screenshot shows the "Local Files" step of the Oracle Database Actions process. The top navigation bar includes "ORACLE Database Actions", "Local Files", and "ADMIN". The main content area displays a list of six data load jobs, each with source and target details and edit/caneta icons:

Source	Target
champs.csv (1K)	CHAMPS
matches.csv (11M)	MATCHES
teambars.csv (25M)	TEAMBARS
participants.csv (90M)	PARTICIPANTS
stats2.csv (249M)	STATS2
stats1.csv (299M)	STATS1

Algumas alterações serão feitas, automaticamente o sistema de carga de dados já gera para o usuário todas as informações necessárias para a carga, como criação das tabelas no banco e os “inserts” dos dados, será colocado antes do nome das tabelas o prefixo “DOTA_”.

Para fazer essas alterações, basta clicar no símbolo de caneta e alterar os nomes das tabelas de destino, conforme mencionado acima.

Sempre observe o nome das colunas e fique atento para que elas não tenham o nome de alguma palavra reservada do Banco Autonomous, como por exemplo a palavra Level e a palavra Cluster são reservadas, necessitando mudar o nome dessas colunas no momento da carga.

Para iniciar a carga dos dados, basta clicar no botão verde.

The screenshot shows the Oracle Database Actions Local Files interface. It displays four separate data load actions, each consisting of a source file and a target table. The actions are:

- Source: chat.csv (50M) Target: DOTA_CHAT
- Source: hero_names.csv (4K) Target: HERO_NAMES
- Source: match.csv (3M) Target: DOTA_MATCH
- Source: players.csv (124M) Target: DOTA_PLAYERS

The screenshot shows the Oracle Database Actions Local Files interface again, but this time the status bar at the top indicates "Status: Running (0/4) - Total time 00:05 seconds". The four data load actions are listed as follows:

- Source: chat.csv (50M) Target: DOTA_CHAT (status: 00:05)
- Source: hero_names.csv (4K) Target: HERO_NAMES
- Source: match.csv (3M) Target: DOTA_MATCH
- Source: players.csv (124M) Target: DOTA_PLAYERS

Após a carga completa é possível ver os detalhes da carga clicando no símbolo de informações.

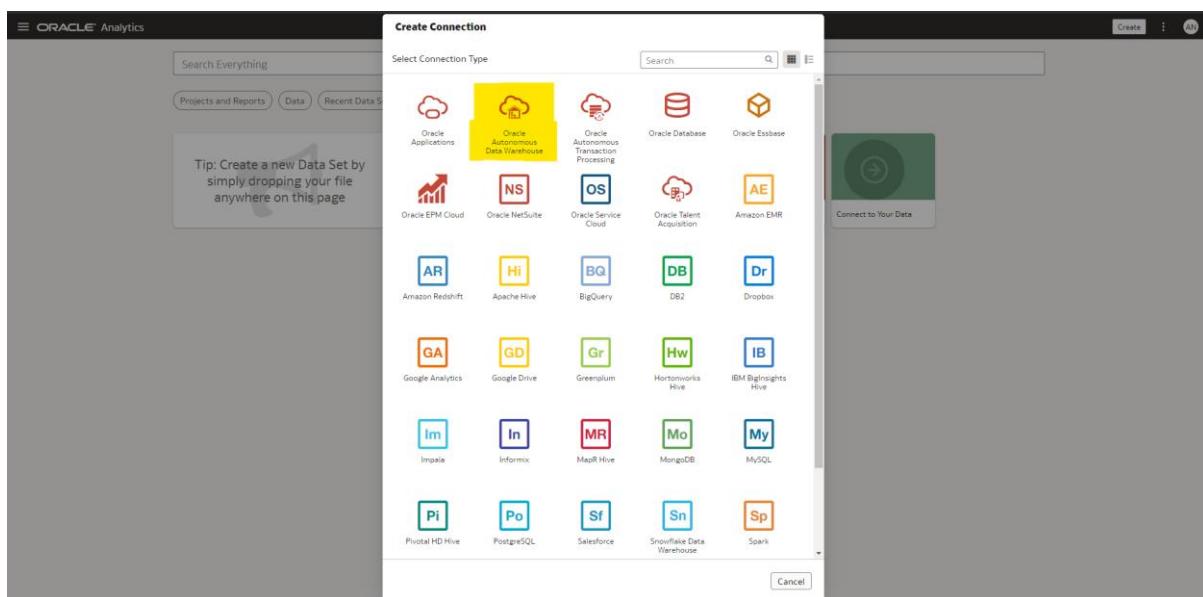
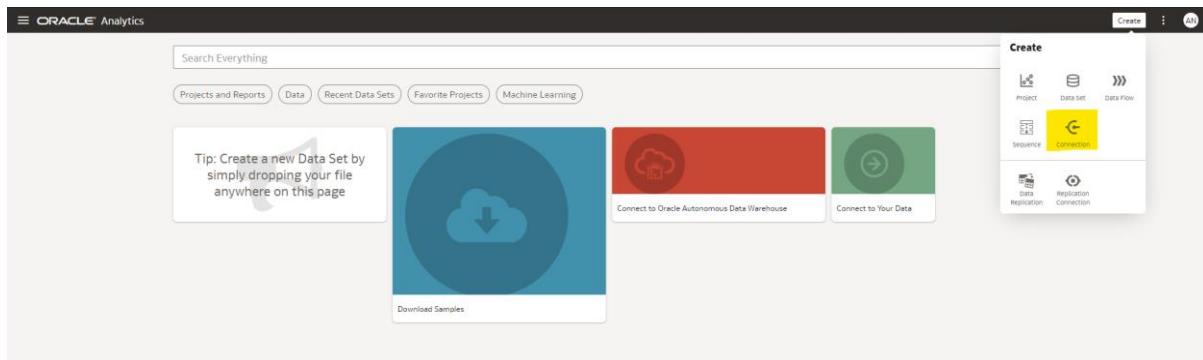
The screenshot shows the Oracle Database Actions Local Files interface with the status bar indicating "Status: Completed (4/4) - Total time 03:56 seconds". All four data load actions are now marked with green checkmarks, indicating successful completion. The actions are listed as follows:

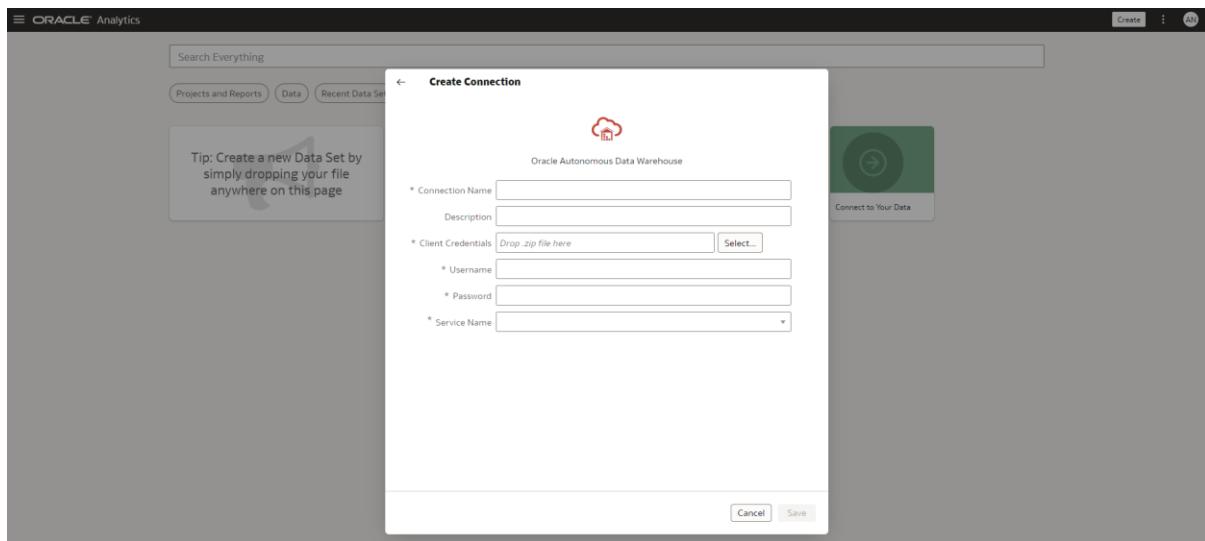
- Source: chat.csv (50M) Target: DOTA_CHAT (status: ✓)
- Source: hero_names.csv (4K) Target: HERO_NAMES (status: ✓)
- Source: match.csv (3M) Target: DOTA_MATCH (status: ✓)
- Source: players.csv (124M) Target: DOTA_PLAYERS (status: ✓)

In the bottom right corner, there are two buttons: "Browse Catalog" and "Done".

Conectando o OAC ao ADW

Vamos até a página inicial do OAC para conectar ele ao ADW, clique em Criar e depois em Conexão e selecione Oracle Autonomous Data Warehouse.





Mas antes de criar a conexão é preciso fazer o Download da Wallet do ADW, conforme abaixo.

Para conectar com o ADW precisamos fazer o download do wallet na página do console na OCI.

Menu: ADW > Clique em DB Connection > e depois em download Wallet.

Depois de cadastrar uma senha e baixar o arquivo .zip do wallet, ela será usada dentro do Oracle Analytics conforme a imagem abaixo.

[←](#) **Create Connection**

 Oracle Autonomous Data Warehouse

* Connection Name

Description

* Client Credentials [Select...](#)

* Username

* Password

* Service Name

[Cancel](#) [Save](#)

Com a conexão com o banco feita, a mesma pode ser verificada indo no menu no canto esquerdo da tela e depois na aba de Dados e em Conexões e lá pode ser visto a conexão que acabou de ser criada, feito isso tudo pronto para começar a criar os Conjuntos de dados.

The screenshot shows the Oracle Analytics interface, specifically the 'Data' section under 'Connections'. The 'Connections' tab is selected, highlighted in yellow. A single connection entry is listed:

Type	Name	Connection Type	Description	Replication Connection	Owner	Modified
Cloud	EXEMPLO	Oracle Autonomous Dat...			armando.neto@ora...	Just now

The left sidebar includes links for Home, Catalog, Data (which is selected and highlighted in blue), Machine Learning, Jobs, and Console.

Preparação para a Análise Exploratória

Preparação dos conjuntos de dados

Primeiro será criado o conjunto de dados com as informações dos jogadores.

The image shows two screenshots of the Oracle Analytics interface. The top screenshot displays the main dashboard with a search bar, navigation tabs (Projects and Reports, Data, Recent Data Sets, Favorite Projects, Machine Learning), and a 'Create' sidebar. The 'Create' sidebar includes options for Project, Data Set (highlighted in yellow), Data Flow, Sequence, Connection, Data Replication, and Replication Connection. A tip box suggests creating a new Data Set by dropping a file. Below the dashboard are three buttons: 'Connect to Oracle Autonomous Data Warehouse' (red), 'Download Samples' (blue), and 'Connect to Your Data' (green). The bottom screenshot shows a detailed 'Create Data Set' dialog box. This dialog has tabs for 'From a File, Subject Area, or Connection' and 'Create Connection'. It features three main options: 'Drop data file here or click to browse' (with an upload icon), 'Local Subject Area' (with a database icon), and 'ORACLEDB' (with a cloud icon). A 'Search' bar and a 'Cancel' button are also present. The background of the main window shows a dark sidebar with 'Add Data Set' and 'Results' sections, and a message in the top right corner stating 'There are no recommendations for this data set.'

Untitled Data Set

Search

ORCLADW

- Manual Query
- ADMIN
- APEX_200200
- APPQOSYS
- ARMANDO
 - BANK_ACCOUNTS
 - BANK_TNS
 - CLOUD_INGEST_JOURNAL
 - CLOUD_INGEST_LOG
 - CSGO_MM_MASTER_DEMOS
 - DATA_CHAT
 - DATA_HERO_NAMES
 - DATA_MATCH
 - DATA_MATCH_OUTCOMES
 - DATA_PLAYERS
 - GRAPH_ACCOUNT
 - GRAPH_CUSTOMER
 - GRAPH_MERCHANT
 - GRAPH_PARENT_OF
 - GRAPH_PURCHASED
 - GRAPH_TRANSFER
 - HDI_PREDICT
 - HDI_TRAINING
 - HOTEL_DEMAND
 - INS_AGENCYPERF
 - INS_CHARGES
 - LOG_FUELCONSUMPTION
 - LOL_CHAMPIONS
 - LOL_MATCHES
 - LOL_PARTICIPANTS
 - STARC

Join Diagram

Drag and drop or double click a table to add it to the data set.

Untitled Data Set

Search

ORCLADW

- Manual Query
- ADMIN
- APEX_200200
- APPQOSYS
- ARMANDO
 - BANK_ACCOUNTS
 - BANK_TNS
 - CLOUD_INGEST_JOURNAL
 - CLOUD_INGEST_LOG
 - CSGO_MM_MASTER_DEMOS
 - DATA_CHAT
 - DATA_HERO_NAMES
 - DATA_MATCH
 - DATA_MATCH_OUTCOMES
 - DATA_PLAYERS

HERO_NAME	HERO_ID	HERO_LOCALIZE...
A HERO_NAME	1 113	This column contains 100% unique values.
B HERO_ID		This column contains 100% unique values.
C HERO_LOCAL...		

Join DOTA_PLAYERS - DOTA_HERO_NAMES

DOTA_PLAYERS DOTA_HERO_NAMES

Inner

Select a column Select a column

Add Join Condition

Untitled Data Set

Search

ORCLADW

- Manual Query
- ADMIN
- APEX_200200
- APPQOSYS
- ARMANDO
 - BANK_ACCOUNTS
 - BANK_TNS
 - CLOUD_INGEST_JOURNAL
 - CLOUD_INGEST_LOG
 - CSGO_MM_MASTER_DEMOS
 - DATA_CHAT
 - DATA_HERO_NAMES
 - DATA_MATCH
 - DATA_MATCH_OUTCOMES
 - DATA_PLAYERS
 - GRAPH_ACCOUNT

HERO_NAME	HERO_ID	HERO_LOCALIZE...
A HERO_NAME	1 113	This column contains 100% unique values.
B HERO_ID		This column contains 100% unique values.
C HERO_LOCAL...	1	Anti-Mage

Join DOTA_PLAYERS - DOTA_HERO_NAMES

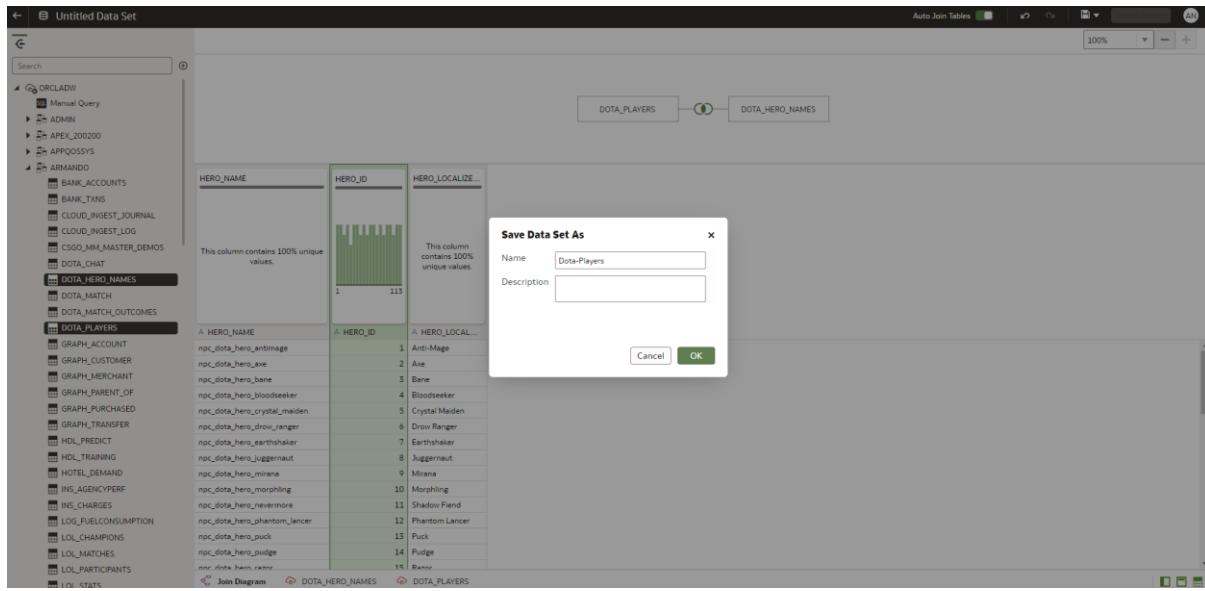
DOTA_PLAYERS DOTA_HERO_NAMES

Inner

HERO_ID HERO_ID

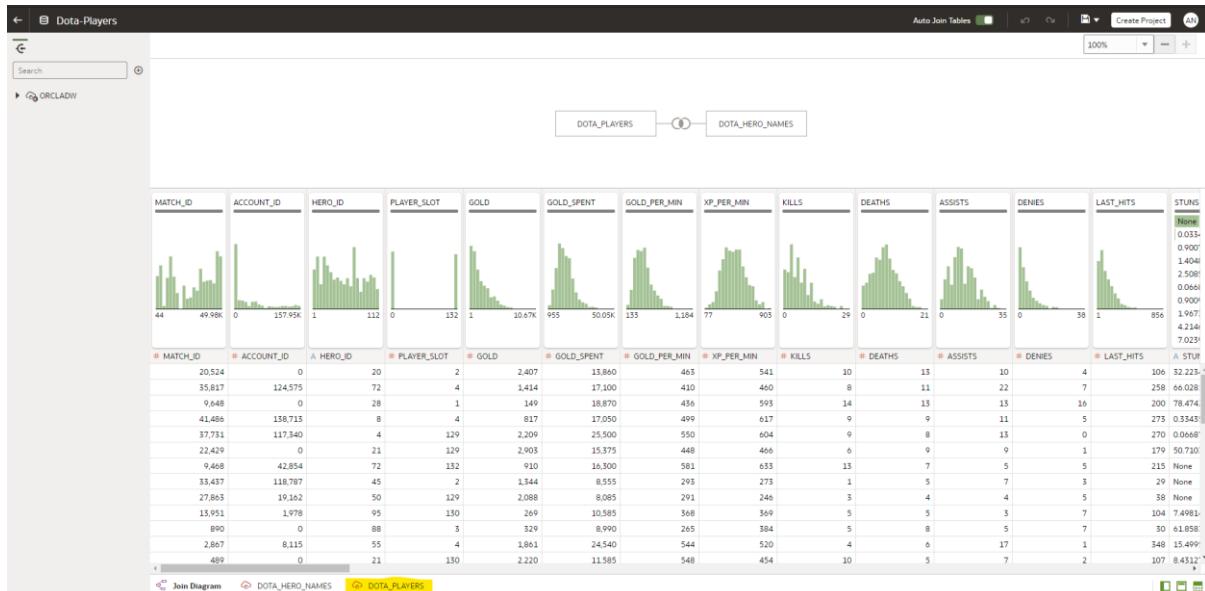
Add Join Condition

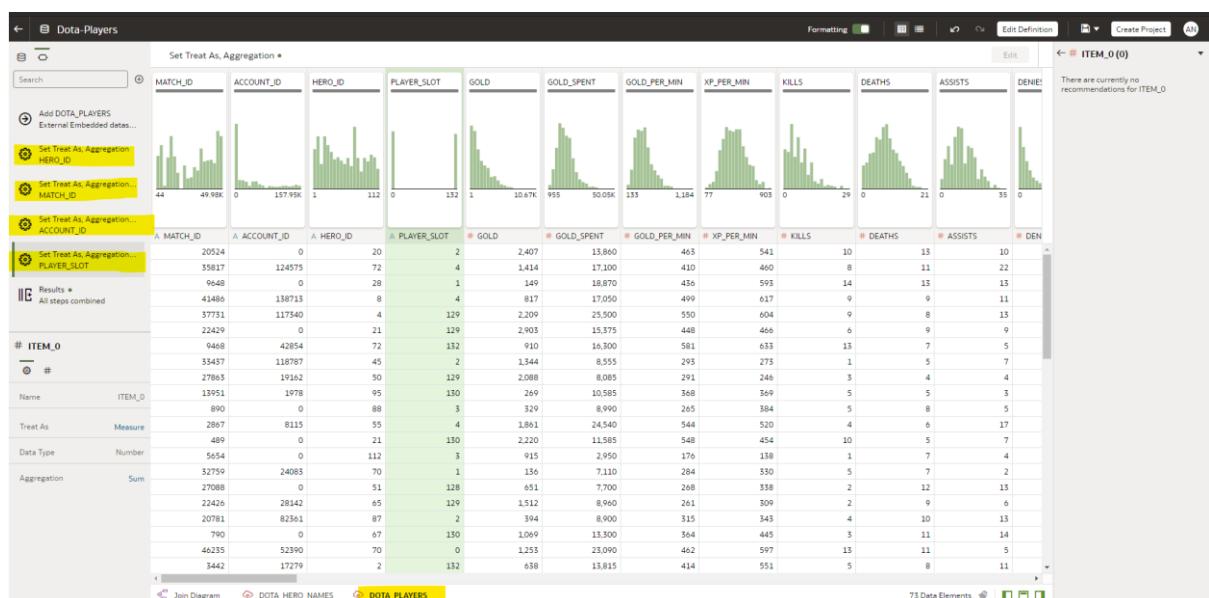
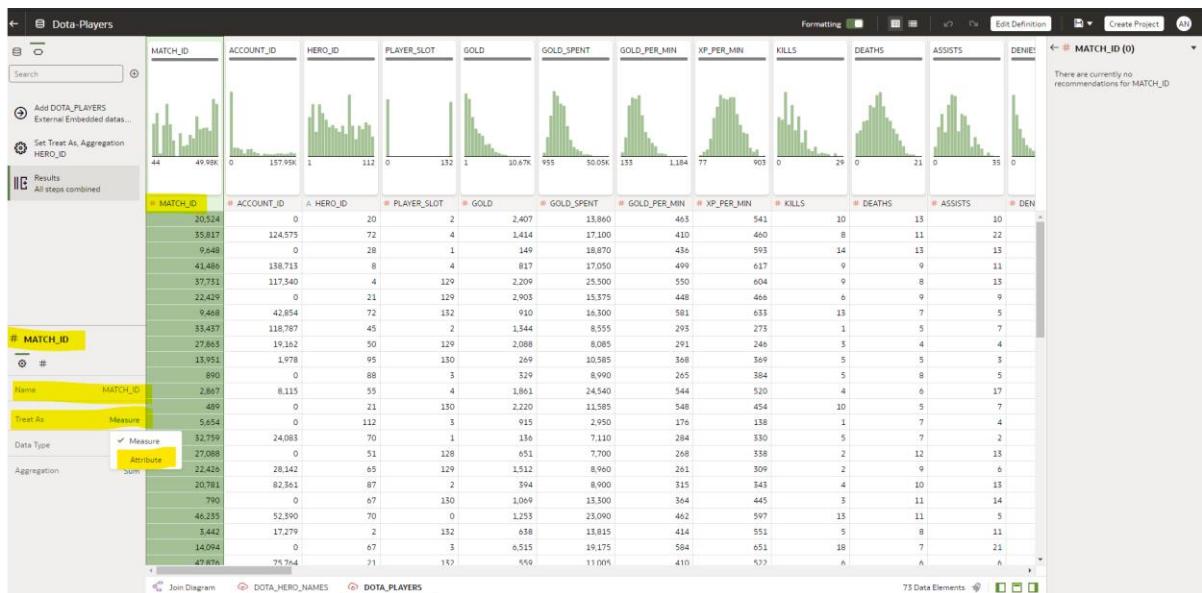
Salve o Conjunto de Dados.



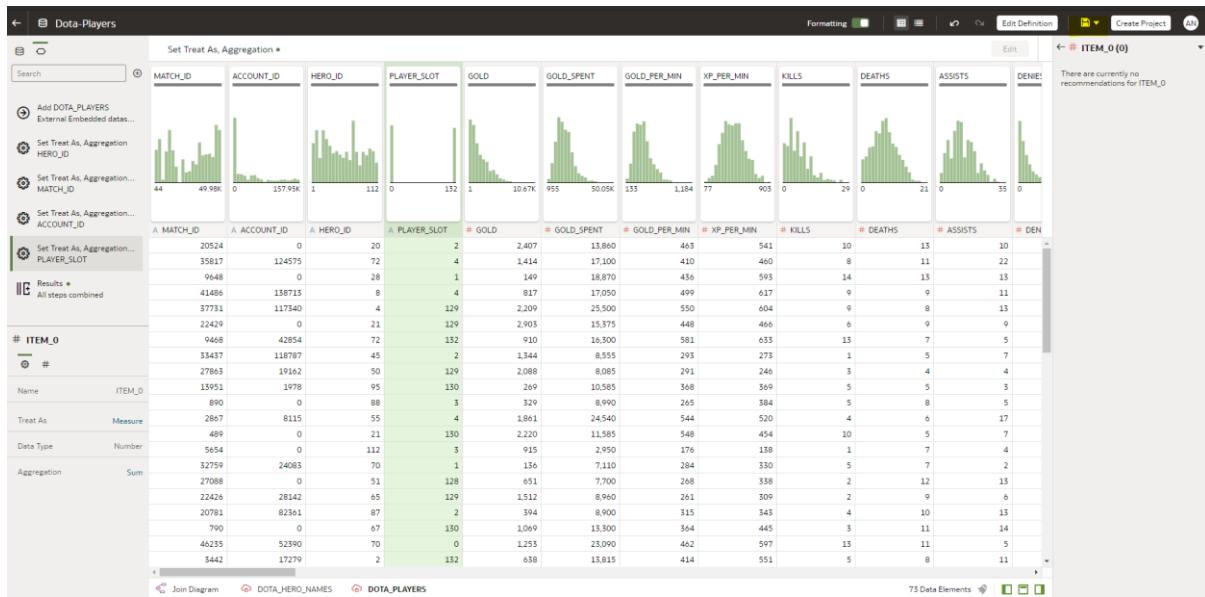
Após criado os Joins, será editado, tabela a tabela, conforme as edições descritas abaixo:

- **DOTA_PLAYERS:** Será alterado como o Analytics trata algumas colunas, basicamente muita das informações estão como números, mas serão utilizadas como atributos, logo é possível alterar como o Analytics interpretará a informação.

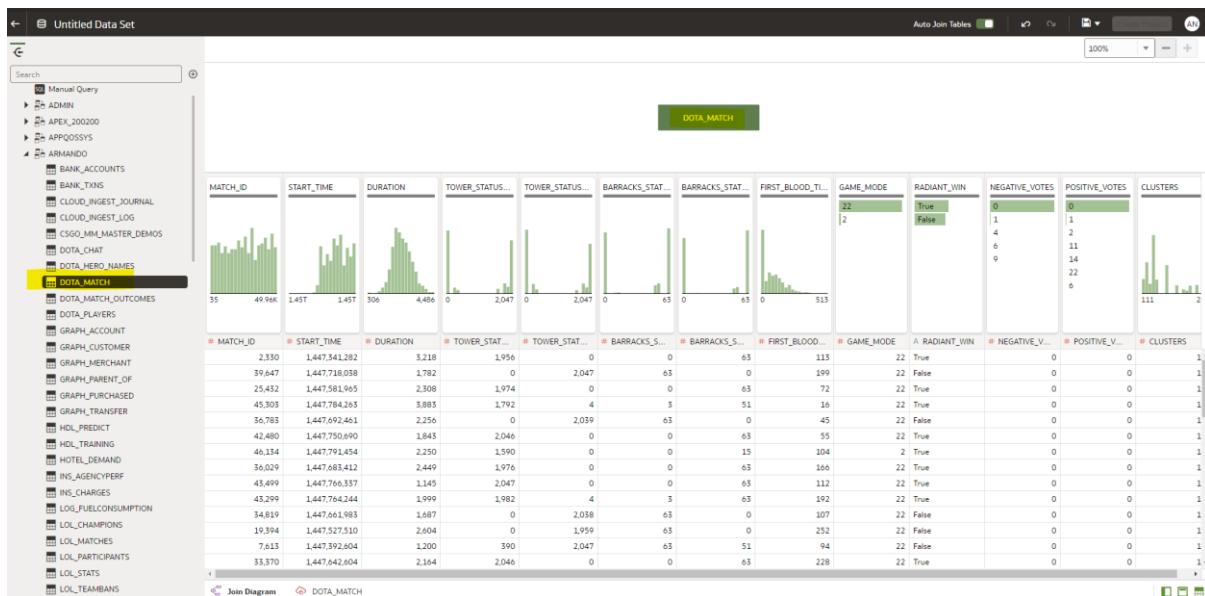




Para salvar os passos que foram feitos, clique no símbolo de salvar no canto superior direito da tela.

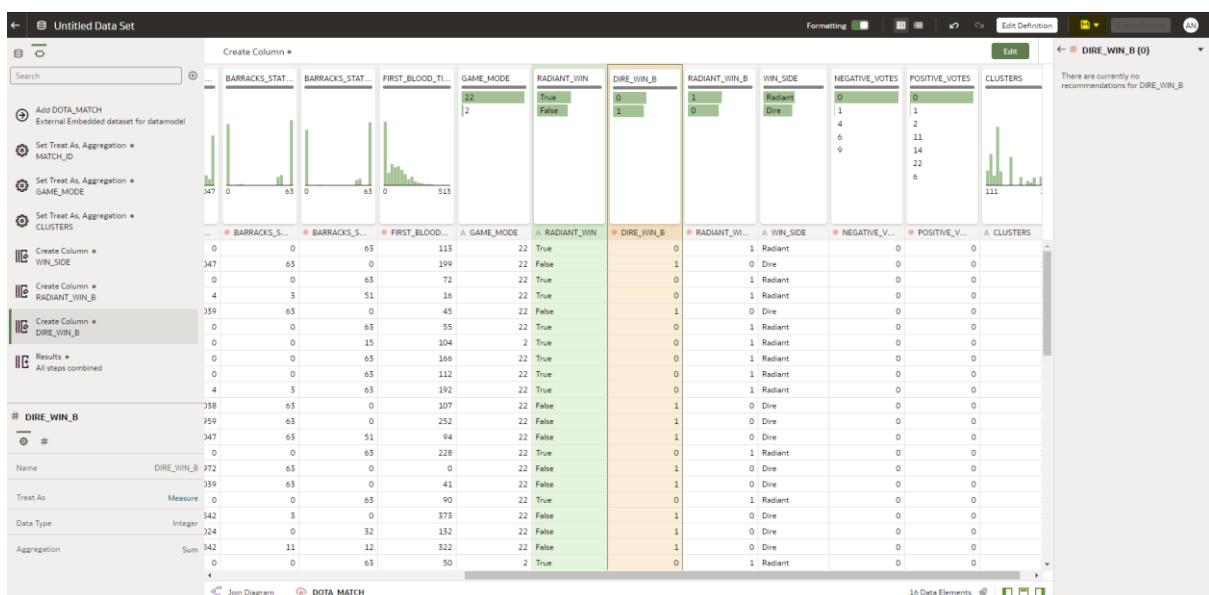
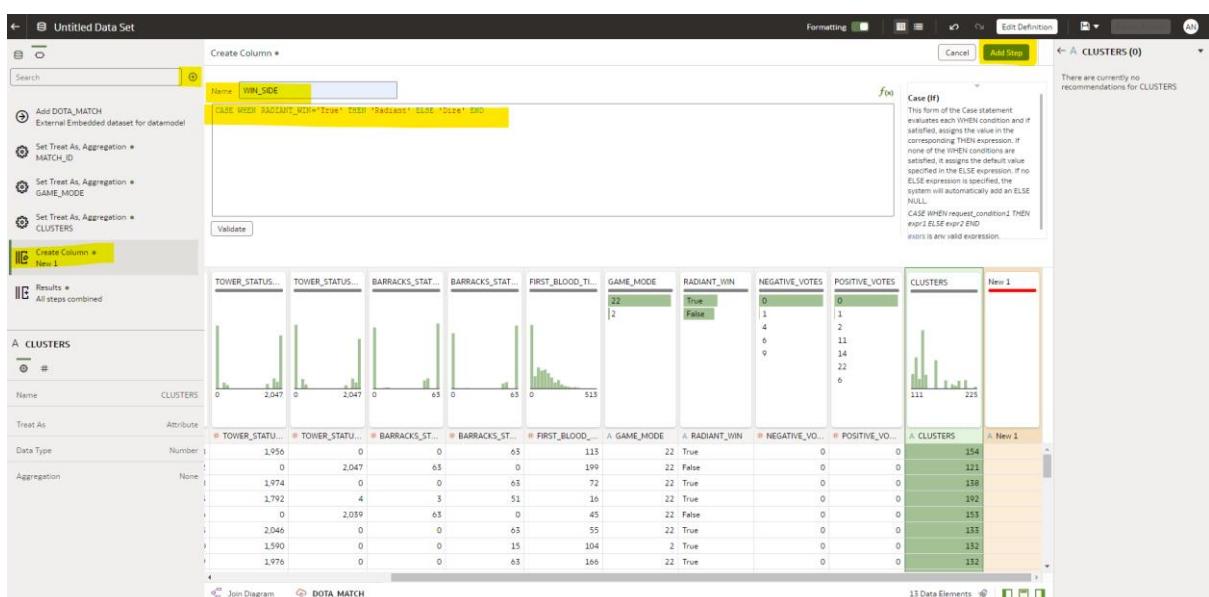
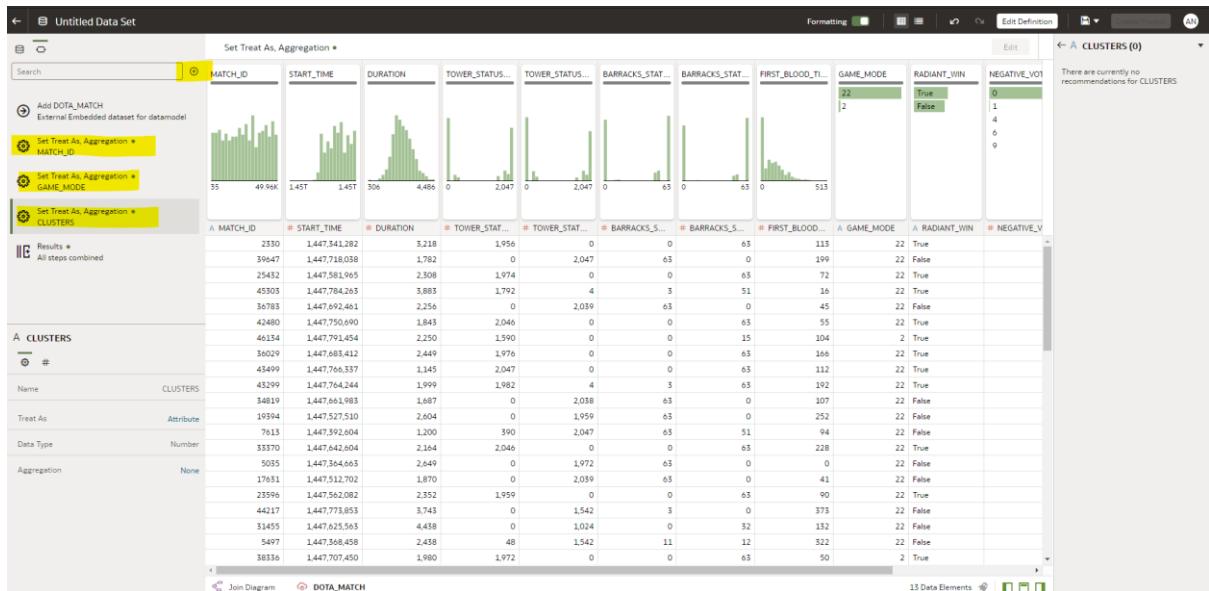


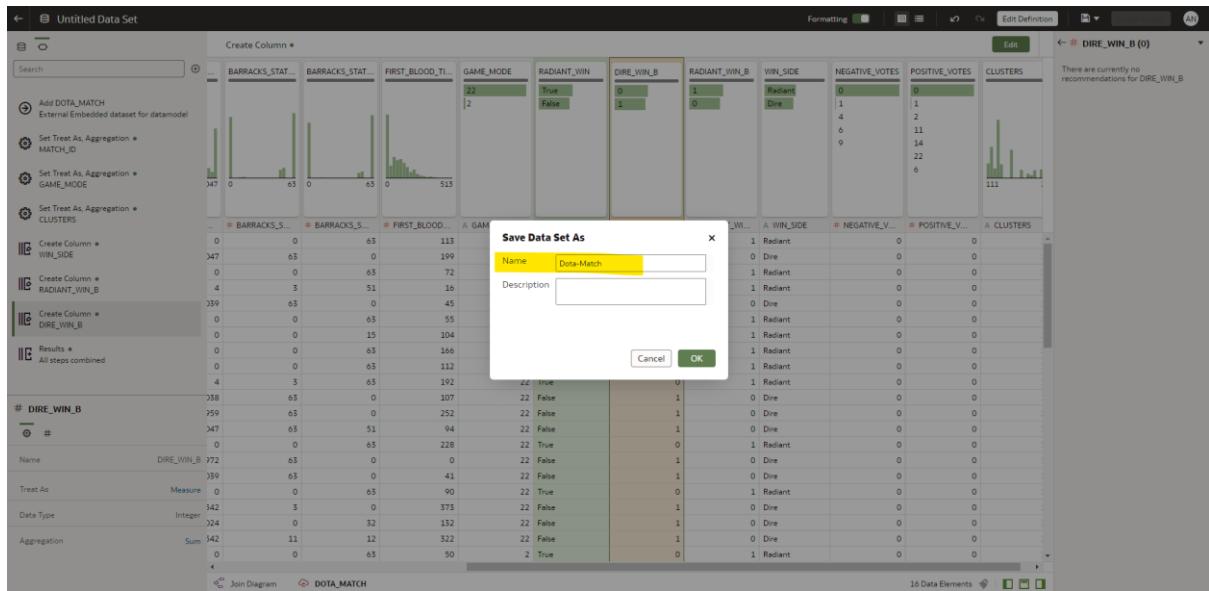
Será criado outro conjunto de dados, o que tem as informações das partidas (DOTA_MATCH), o passo a passo para a criação é o mesmo do já executado neste documento.



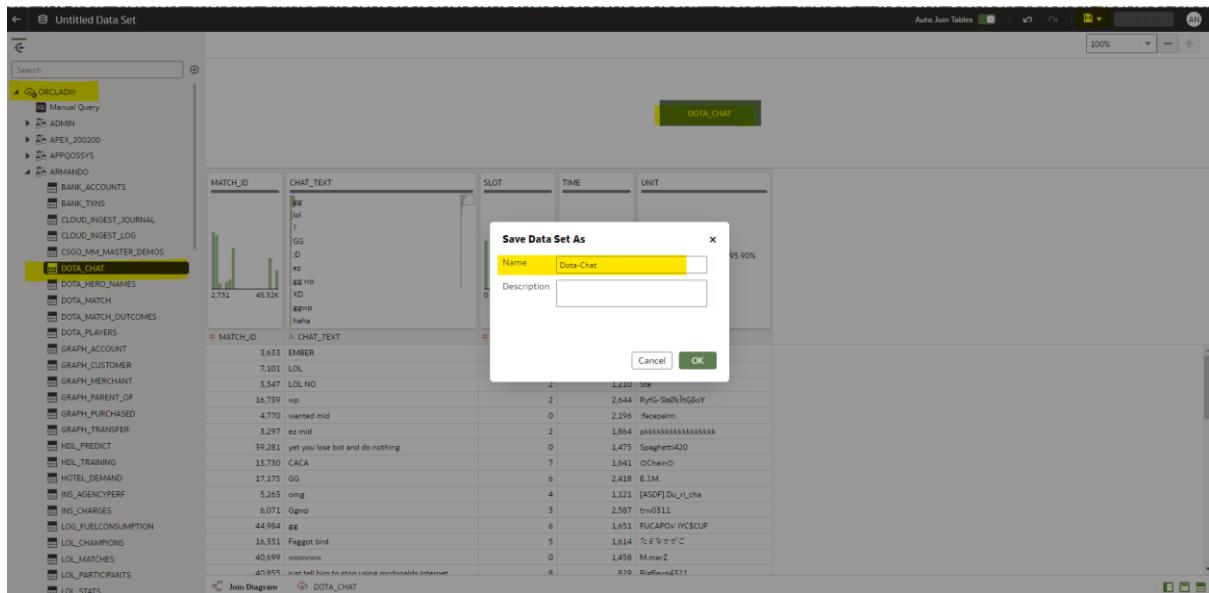
Serão feitas algumas alterações em como o Analytics trata algumas colunas e serão criadas 3 colunas novas, são elas:

- WIN_SIDE = CASE WHEN RADIANT_WIN='True' THEN 'Radiant' ELSE 'Dire' END
- RADIANT_WIN_B = CASE WHEN RADIANT_WIN='True' THEN 1 ELSE 0 END
- DIRE_WIN_B = CASE WHEN RADIANT_WIN='True' THEN 0 ELSE 1 END





Por fim será criando mais um conjunto de dados respectivo a tabela DOTA_CHAT, para gerar a nuvem de palavras na análise do chat no Dota2. O Processo de criação é o mesmo já executado e neste conjunto de dados não será feita nenhuma alteração.



Para que a nuvem de palavras seja possível de ser criada, o conjunto de dados precisa passar por um algoritmo de tokenização das palavras dentro das frases, será usado da sinergia entre o Oracle Analytics com o Oracle Autonomous, o Analytics consegue chamar a função de tokenização do banco Autonomous via uma ferramenta dentro do Analytics chamada de Data Flow. Segue o passo a passo

The screenshot shows the Oracle Analytics interface. At the top, there's a search bar labeled "Search Everything". Below it are navigation tabs: "Projects and Reports", "Data", "Recent Data Sets", "Favorite Projects", and "Machine Learning". On the right side, there's a "Create" button with a yellow outline. A tooltip "Data Set" is shown over the "Data Set" icon. Below the "Create" button are icons for "Project", "Sequence", "Data Replication", and "Connection". The main area contains several cards: "Data-Match" (cloud icon), "Data-Players" (cloud icon), a tip card with text "Tip: Create a new Data Set by simply dropping your file anywhere on this page", a large blue "Download Samples" card with a cloud icon and a download arrow, and a red "Connect to Oracle Autonomous Data Warehouse" card with a cloud icon. At the bottom, there's a section titled "Data Sets" with three cards: "Data-Chat", "Data-Match", and "Data-Players".

The screenshot shows a software interface for data analysis. On the left, there's a sidebar with various menu items such as 'Add Data', 'Join', 'Union Rows', 'Filter', 'Aggregate', 'Save Data Set', 'Create Ensemble Cube', 'Add Columns', 'Select Columns', 'Rename Columns', 'Transform Column', 'Merge Columns', 'Split Columns', 'Bin', 'Group', 'Branch', 'Cumulative Value', 'Time Series Forecast', 'Analyze Sentiment', 'Train Numeric Prediction', 'Train Multi-Classifier', 'Train Clustering', 'Train Binary Classifier', and 'Apply Model'. The main area has a search bar at the top. A message 'No data added.' is displayed above a '+' button and a placeholder text 'Click '+' to add data.' Below this is a 'Create Data Set' button. The central part of the screen is a 'Add Data Set' dialog box. It contains a search bar, three preview icons for datasets ('Data-Chit', 'Data-Match', 'Data-Players'), and 'Cancel' and 'Add' buttons at the bottom.

The screenshot shows the Data-Chat interface. The left sidebar contains a search bar and a list of operations: Add Data, Join, Union Rows, Filter, Aggregate, Save Data Set, Create Essbase Cube, Add Columns, Select Columns, Rename Columns, Transform Column, Merge Columns, Split Columns, Bin, Group, Branch, Cumulative Value, Time Series Forecast, Analyze Sentiment, Train Numeric Prediction, Train Multi-Classification, Train Clustering, Train Binary Classifier, Apply Model, Database Analytics, and Graph Analytics. The main workspace is titled "Add Data - Data-Chat" and shows a table with columns: MATCH_ID, CHAT_TEXT, SLOT, TIME, and UNIT. The table contains several rows of data.

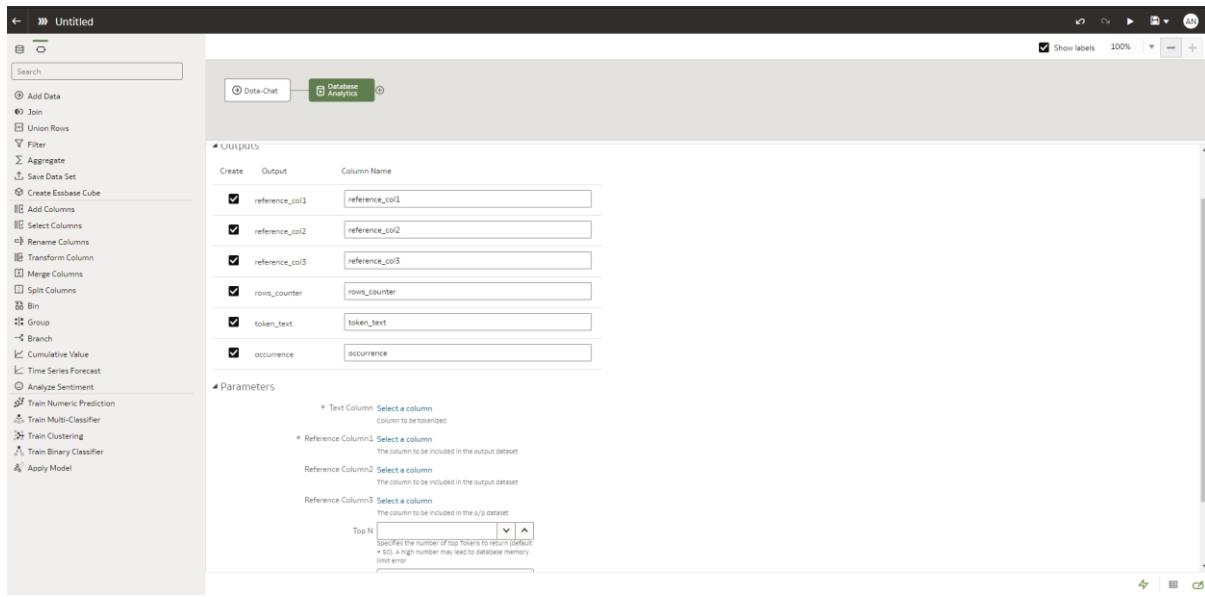
	MATCH_ID	CHAT_TEXT	SLOT	TIME	UNIT
15919	lucky	3	638	wacks.canvas	
15919	lol	6	639	c r a y o n	
15919	?	6	643	c r a y o n	
15919	LUCKY ALWAYS INV	3	645	wacks.canvas	
15919	STUPID DATA	3	647	wacks.canvas	
15919	lucky wr	3	661	wacks.canvas	
15919	RUN	3	777	wacks.canvas	
15919	tang ina	3	1024	wacks.canvas	
15919	Anpo un	3	1025	wacks.canvas	
15919	d man Ing pinaglawenf	3	1035	wacks.canvas	
15919	NC COMBOS	3	1155	wacks.canvas	
15919	STUPID TEAM	3	1161	wacks.canvas	

As ações de Database Analytics e Graph Analytics só aparecem quando o conjunto de dados criado vem de uma conexão com um Banco Oracle ou Autonomous

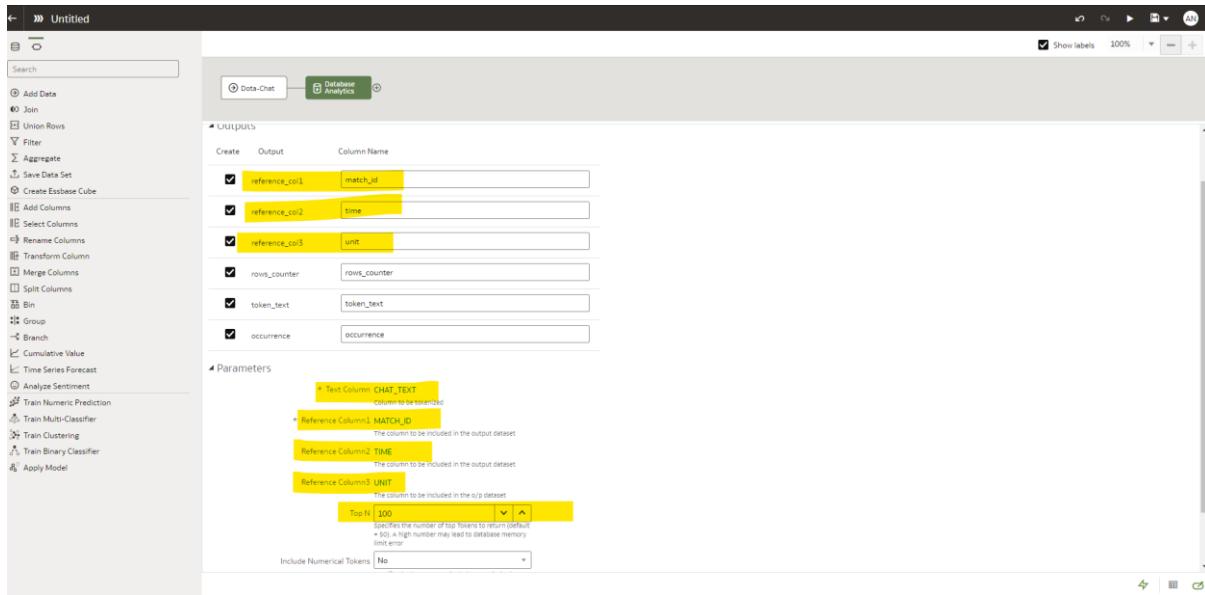
The screenshot shows the Oracle Analytics interface with the title bar "Untitled". On the left, there's a sidebar with various data manipulation and machine learning steps like "Add Data", "Join", "Union Rows", etc. The main workspace has a toolbar at the top with icons for "Add Data", "Join", "Union Rows", "Filter", "Aggregate", "Save Data Set", "Create Essbase Cube", "Add Columns", "Select Columns", "Rename Columns", "Transform Column", "Merge Columns", "Split Columns", "Bin", "Group", "Branch", "Cumulative Value", "Time Series Forecast", "Analyze Sentiment", "Train Numeric Prediction", "Train Multi-Classifier", "Train Clustering", "Train Binary Classifier", "Apply Model", "Database Analytics", and "Graph Analytics". The "Database Analytics" icon is highlighted with a yellow box. Below the toolbar is a table with columns: MATCH_ID, CHAT_TEXT, SLOT, TIME, and UNIT. The table contains several rows of data. At the bottom right of the table, there are checkboxes for "TIME" and "UNIT".

This screenshot shows the "Select Database Analytics" dialog box. It lists various analytics operations: Text Tokenization, Dynamic Clustering, Frequent Itemsets, Time Series, Dynamic Anomaly Detection, Sampling Data, and Un-pivoting Data. The "Text Tokenization" option is highlighted with a yellow box. The background shows the same Oracle Analytics interface as the previous screenshot, with the "Database Analytics" step selected in the toolbar.

Com isso o Oracle Analytics consegue chamar funções do Banco Oracle com a interface gráfica dele, sem necessitar criar nenhum código ou script em SQL ou PL/SQL (Linguagem Scriptural do Banco Oracle). É necessário passar os parâmetros da função.



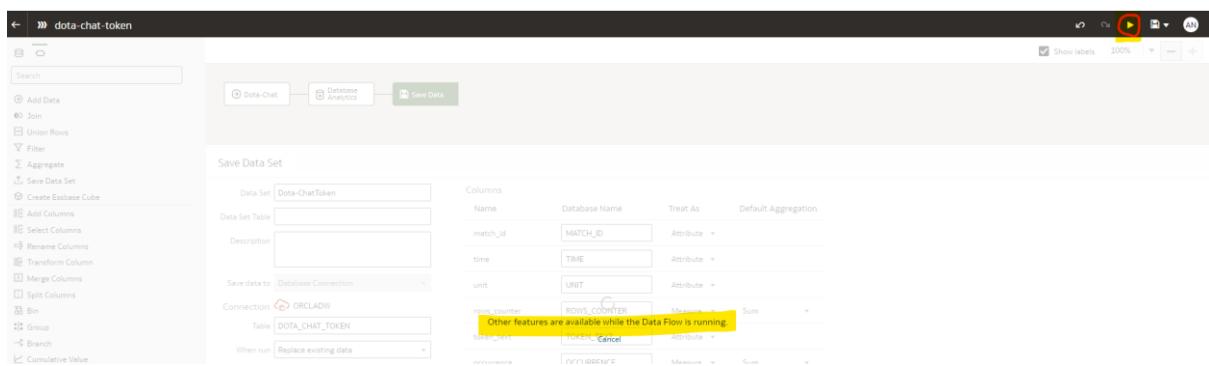
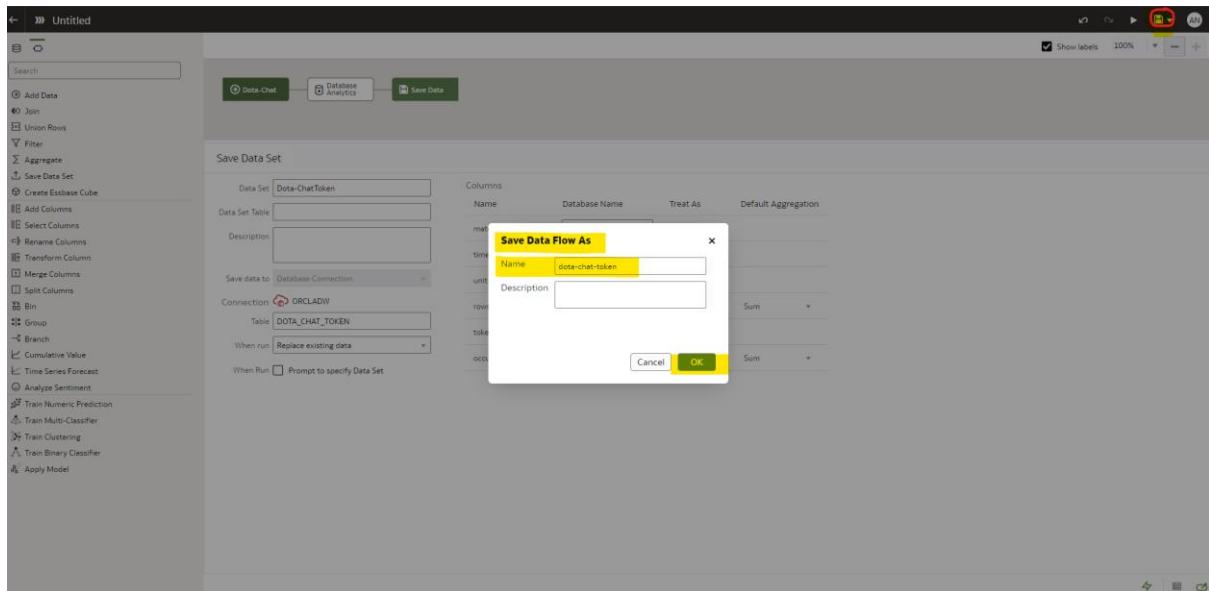
As colunas referenciadas serão a MATCH_ID, UNIT e TIME as colunas de rows_counter, token_text e occurrence são as saídas do script, a coluna que será o texto a ser tokenizado é a CHAT_TEXT e no parâmetro Top N será colocado 100 para vermos as 100 palavras mais enviadas no chat do Dota2, o número é 100 por causa da limitação da visualização de nuvem de palavras do Oracle Analytics, os outros parâmetros serão usados os valores padrão.



The screenshot shows the Microsoft Data Flow interface. The top part displays a flow from 'Data-Chat' to 'Database Analytics'. The 'Database Analytics' stage includes various operations like 'Add Data', 'Join', 'Union Rows', 'Filter', 'Aggregate', 'Save Data', and 'Create Essbase Cube'. The 'Save Data' button is highlighted. Below this, the 'Outputs' section lists columns: 'reference_col1' (match_id), 'reference_col2' (time), 'reference_col3' (unit), 'rows_counter' (rows_id), 'token_text' (token_id), and 'occurrence' (occurrence). The 'Parameters' section specifies 'Text Column: CHAT_TEXT' (Column to be tokenized), 'Reference Column1: MATCH_ID' (The column to be included in the output dataset), 'Reference Column2: TIME' (The column to be included in the output dataset), 'Reference Column3: UNIT' (The column to be included in the o/p dataset), 'Top N: 100' (Specifies the number of top tokens to return (default = 50). A high number may lead to database memory limit error), and 'Include Numerical Tokens: No'. The bottom part shows the 'Save Data Set' configuration window, where the 'Data Set' is named 'DATA_CHAT_TOKEN' and saved to the 'ORCLADW' database connection.

Como está sendo usado o banco de dados para suportar esta análise, a tabela de saída deve ser salva no banco de dados, neste passo é necessário somente dar um nome para o conjunto de dados e a tabela do banco.

Feito isso, salve o Data Flow e execute conforme a imagem abaixo.



A preparação dos dados foi finalizada, agora começará a Análise Exploratória.

Executando a Análise Exploratória

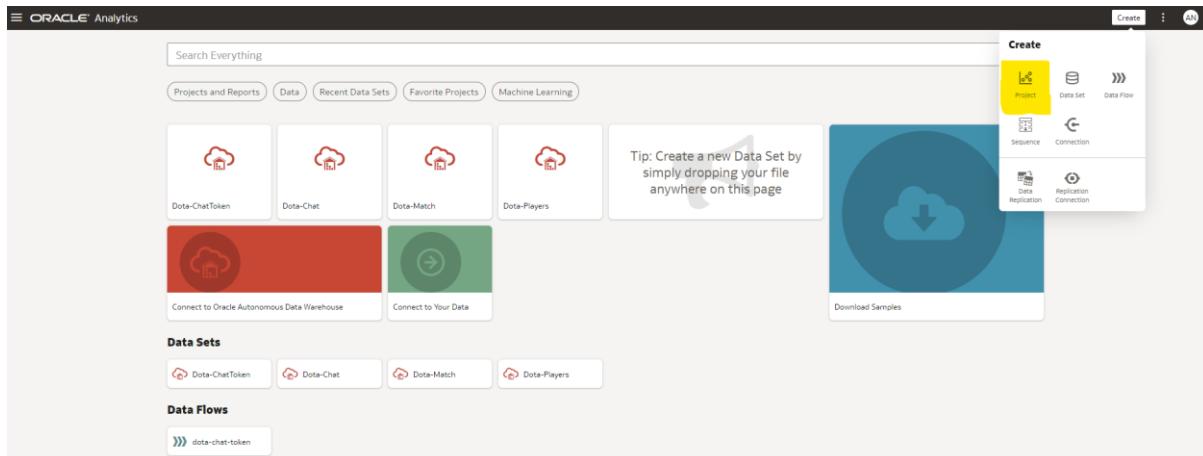
Na análise exploratória a ideia é responder algumas perguntas para a melhorar o entendimento sobre o jogo, o comportamento dos jogadores durante as partidas ranqueadas e destrinchar quais são os melhores campeões, as melhores composições, o que deve ser focado e muito mais, por consequência melhorar com Dados suas habilidades no Dota2.

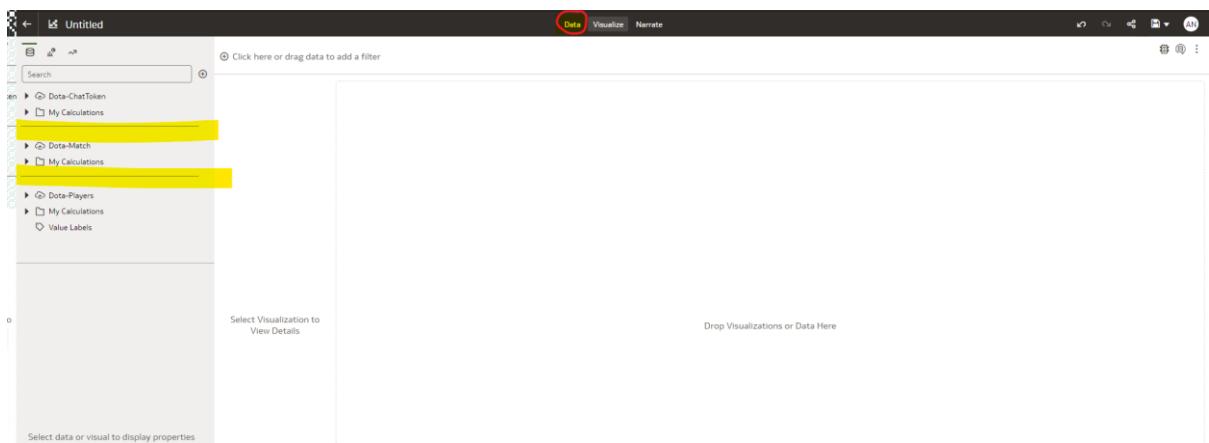
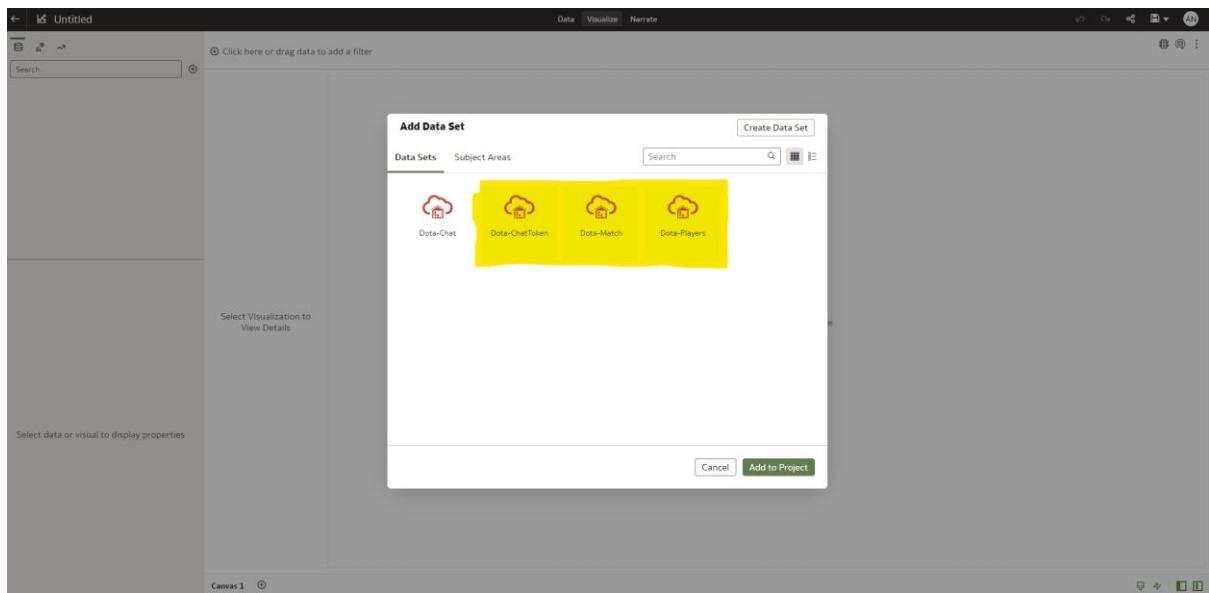
O objetivo da análise é responder algumas perguntas, tais como:

- Visão Geral das Partidas
- Performance dos Jogadores
- Vitórias por Herói
- Análise do Chat

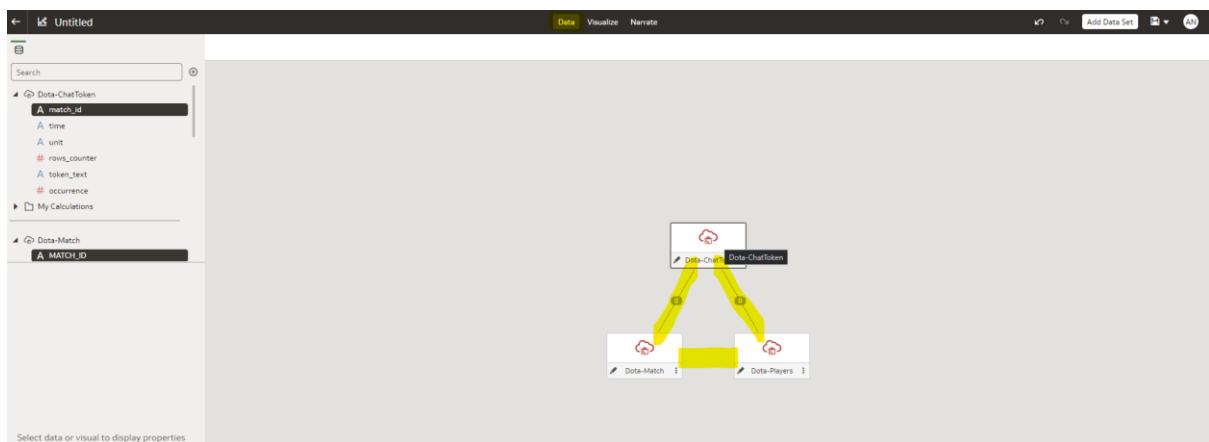
Visão Geral

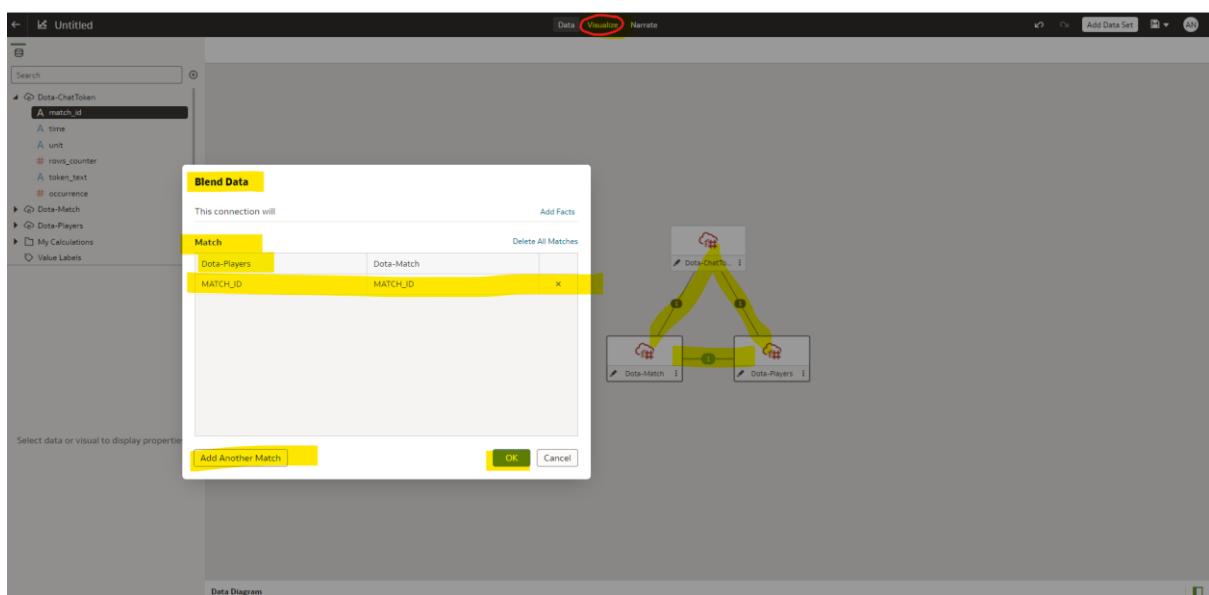
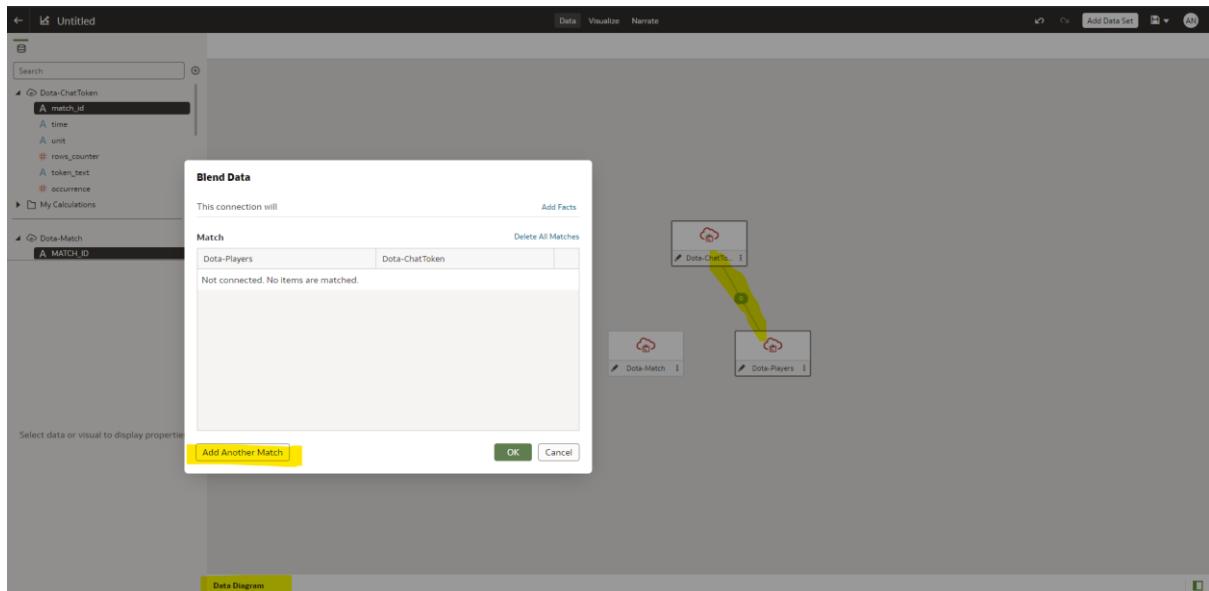
Começando a criação do projeto.





As barras grifadas na imagem acima indicam que os conjuntos de dados adicionados acima eles ainda não estão “se conversando”, não estão com um Join feito, para criar esse Join manualmente vá na aba Data.

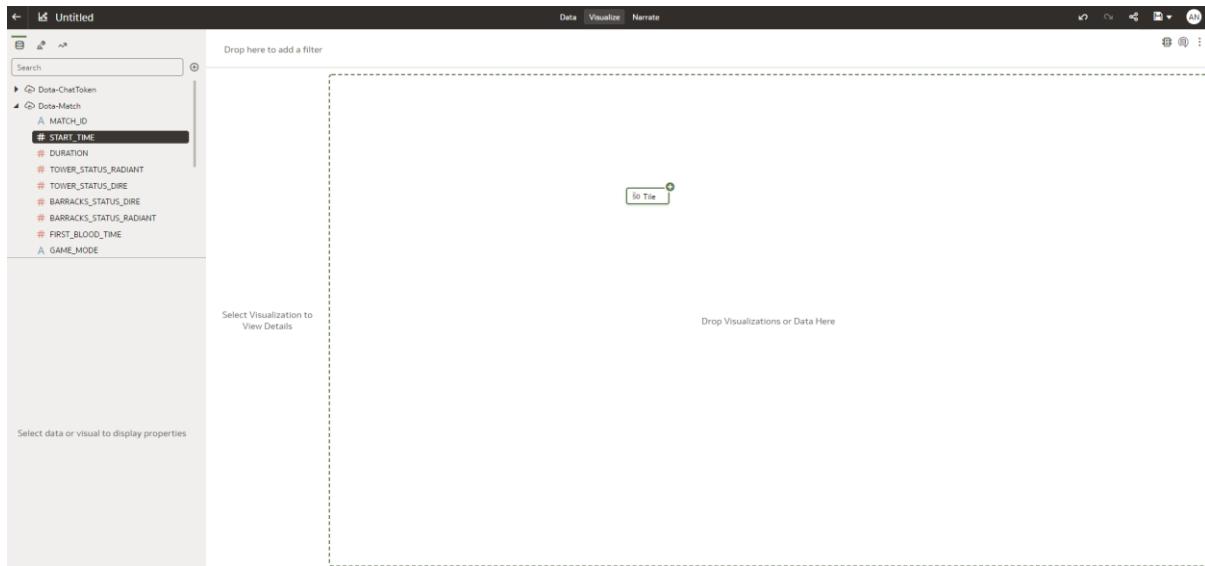




A coluna usada para fazer o Join entre todos os conjuntos de dados é somente a MATCH_ID.

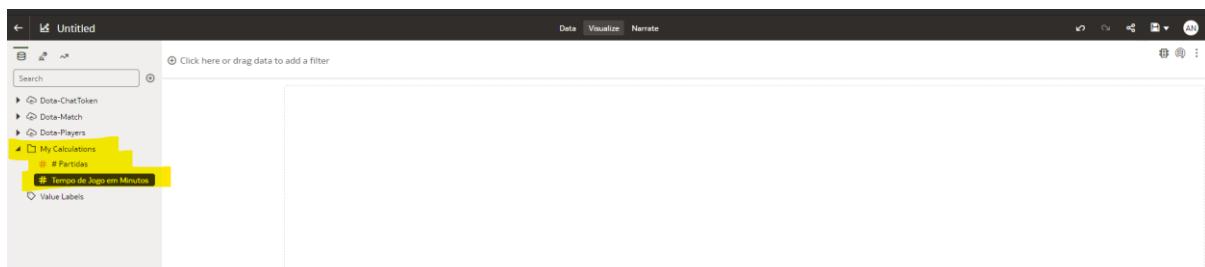
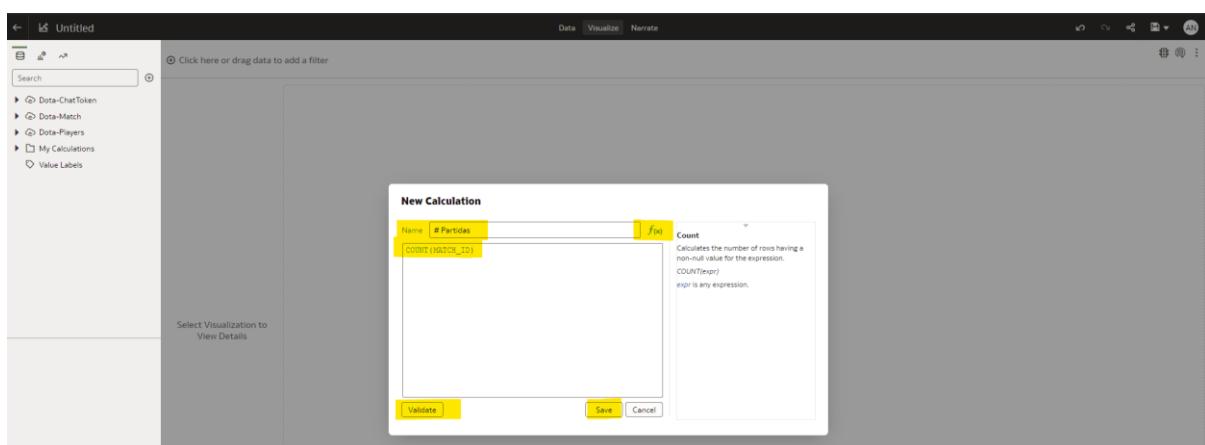
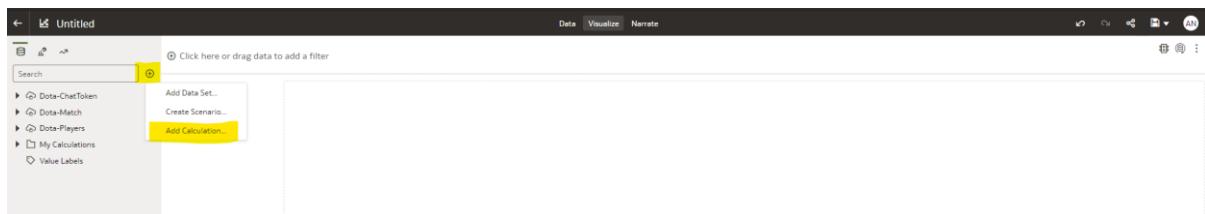
Pronto agora será possível correlacionar as informações dentro desses conjuntos de dados que estavam separados. Volte para a aba de Visualize para começar a análise exploratória.

Pegando e arrastando uma métrica para o campo que está escrito “Soltar Visualizações ou Dados Aqui”, faz com que o OAC crie automaticamente a visualização que ele acha melhor para interpretar o dado, conforme imagem abaixo.

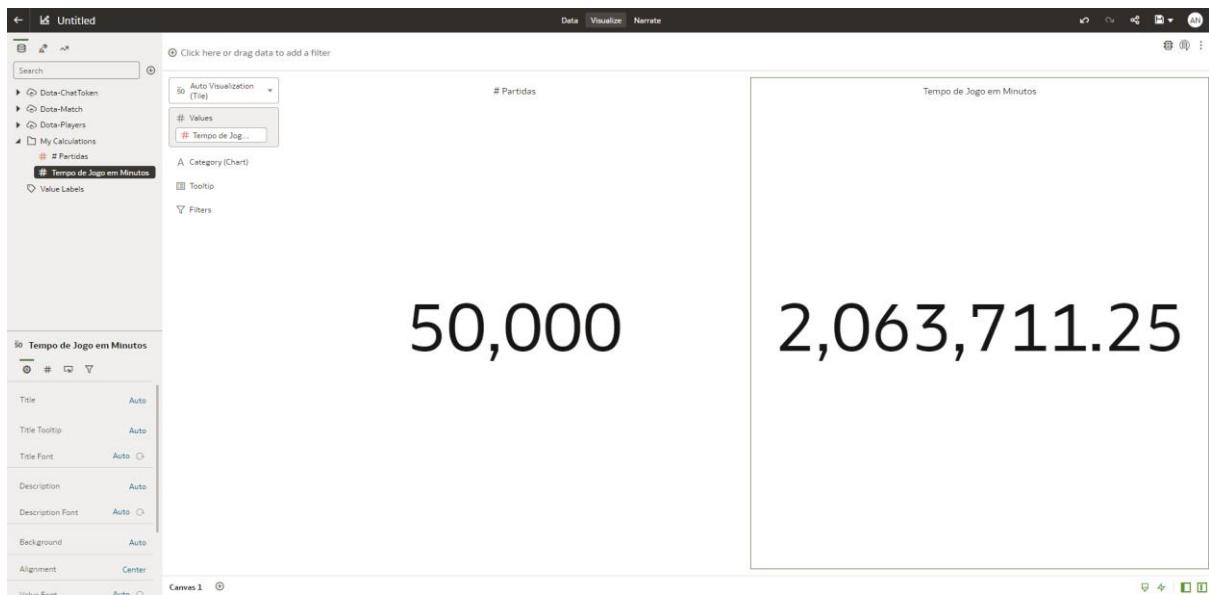
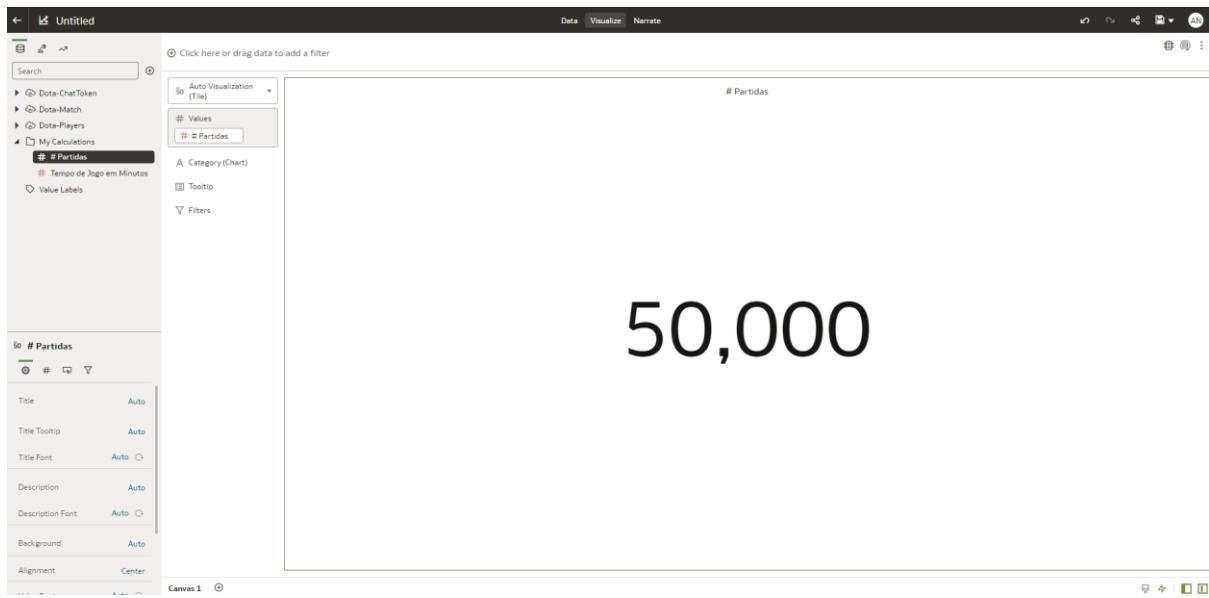


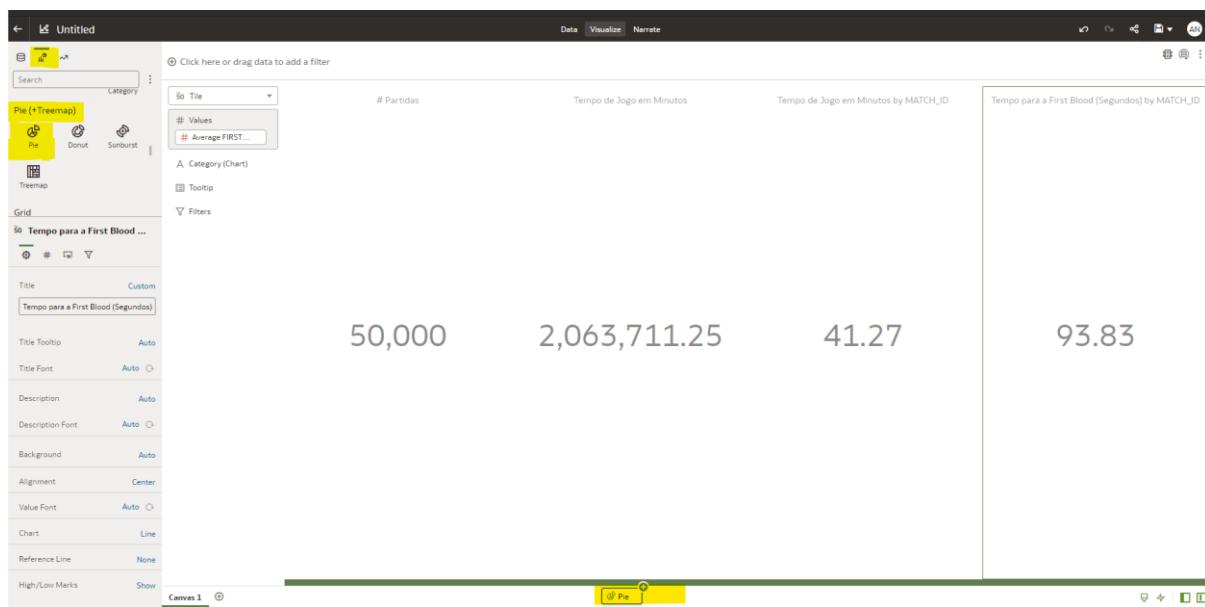
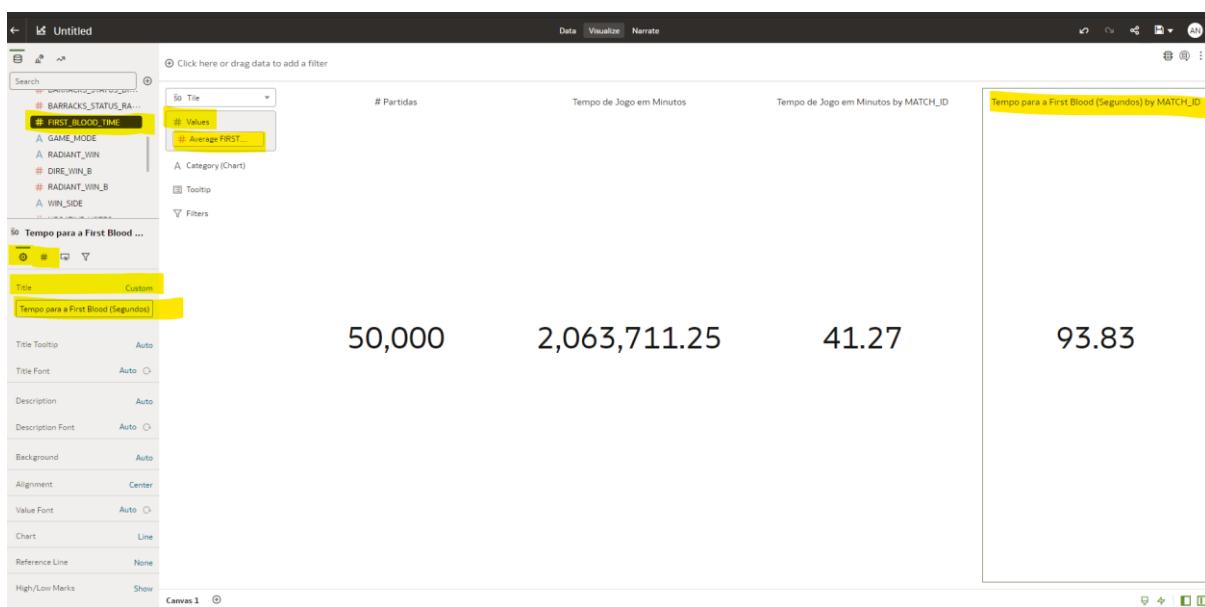
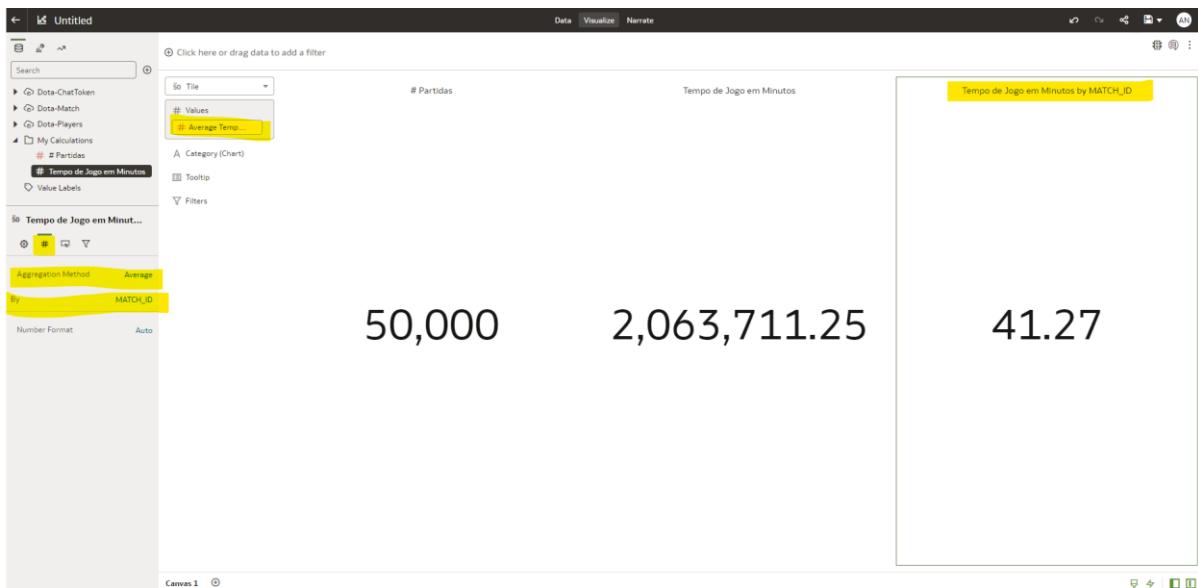
Antes de começar a criação do dashboard serão criadas algumas métricas calculadas para o projeto. São elas:

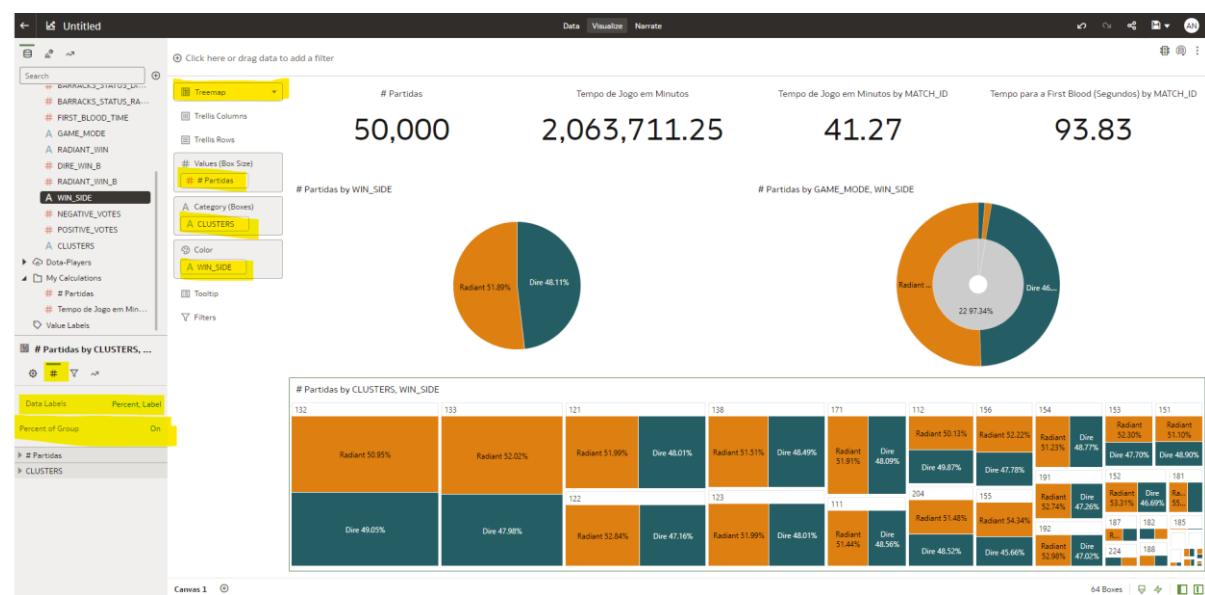
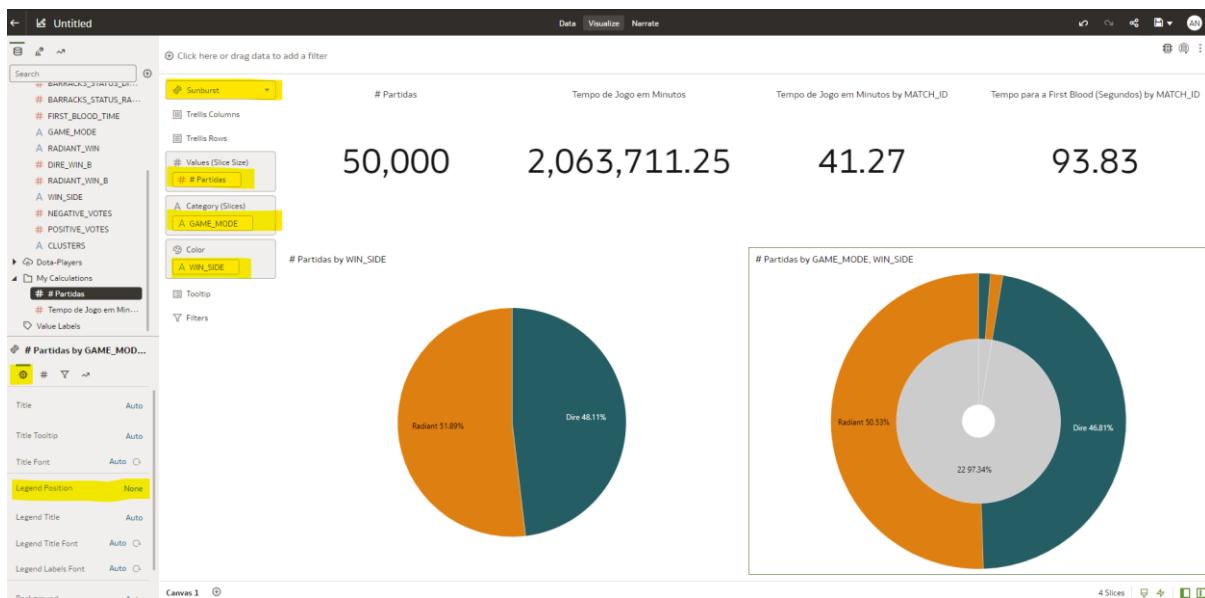
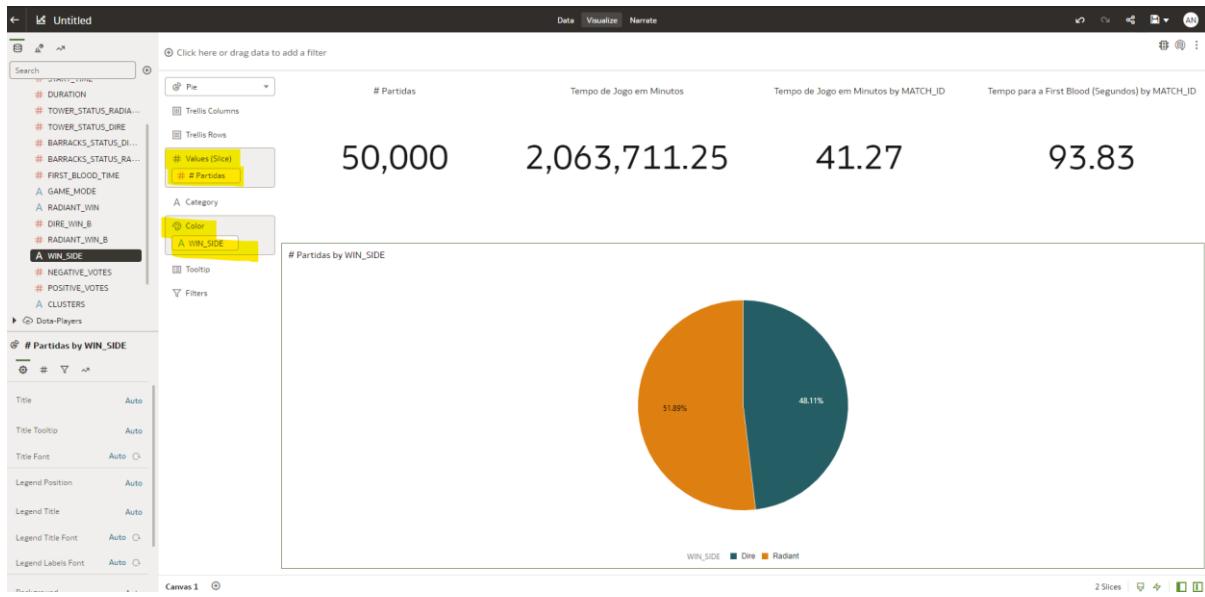
- # Partidas = COUNT(MATCH_ID)
- Tempo de Jogo em Minutos = DURATION/60

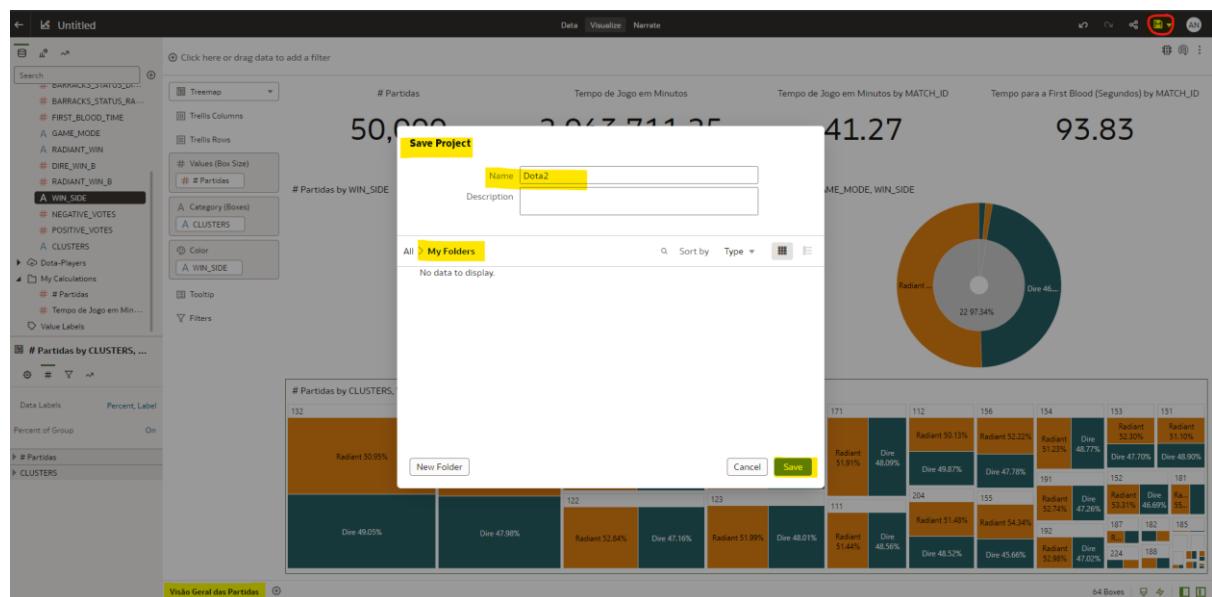
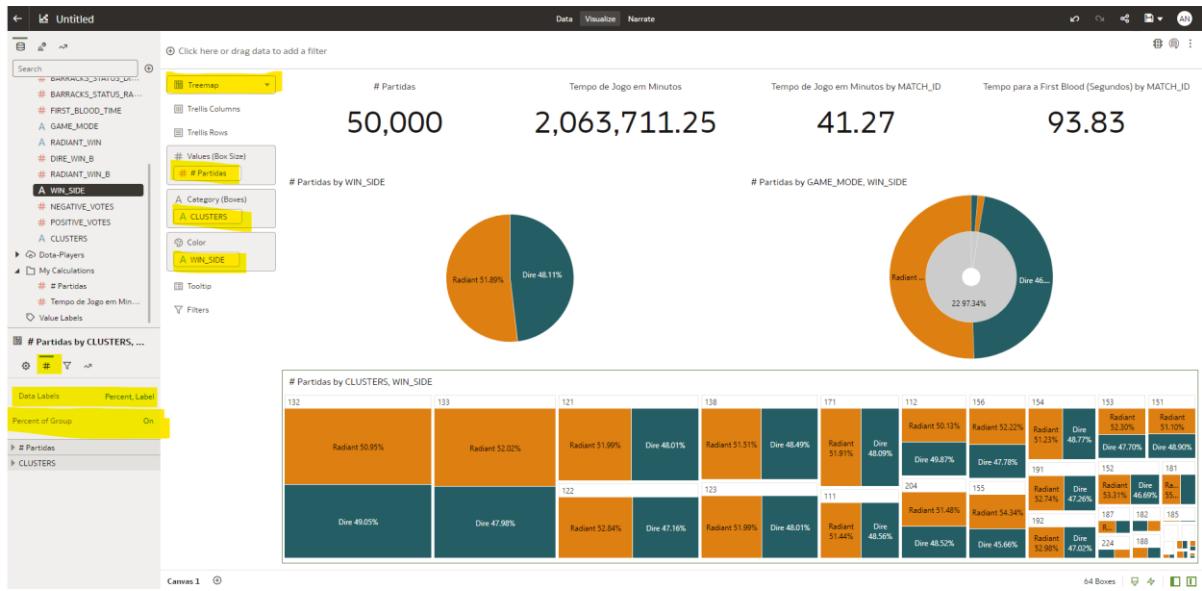


Agora será feita uma série de imagens para mostrar o passo a passo da construção do Dashboard de Visão Geral e ao final será feito comentários e insights sobre os dados.









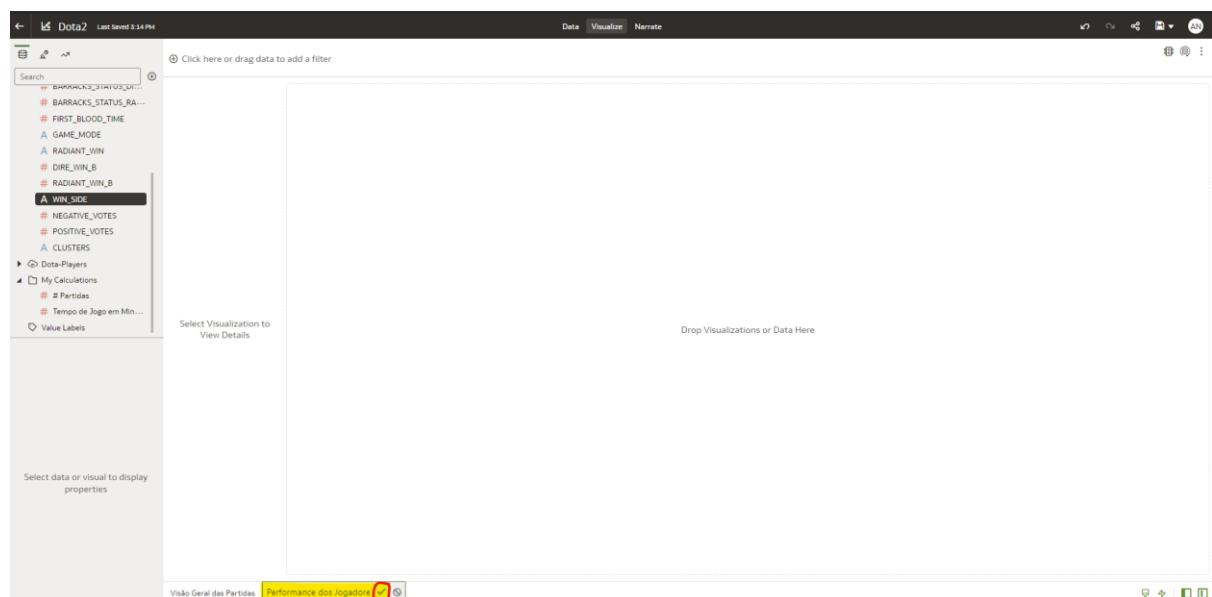
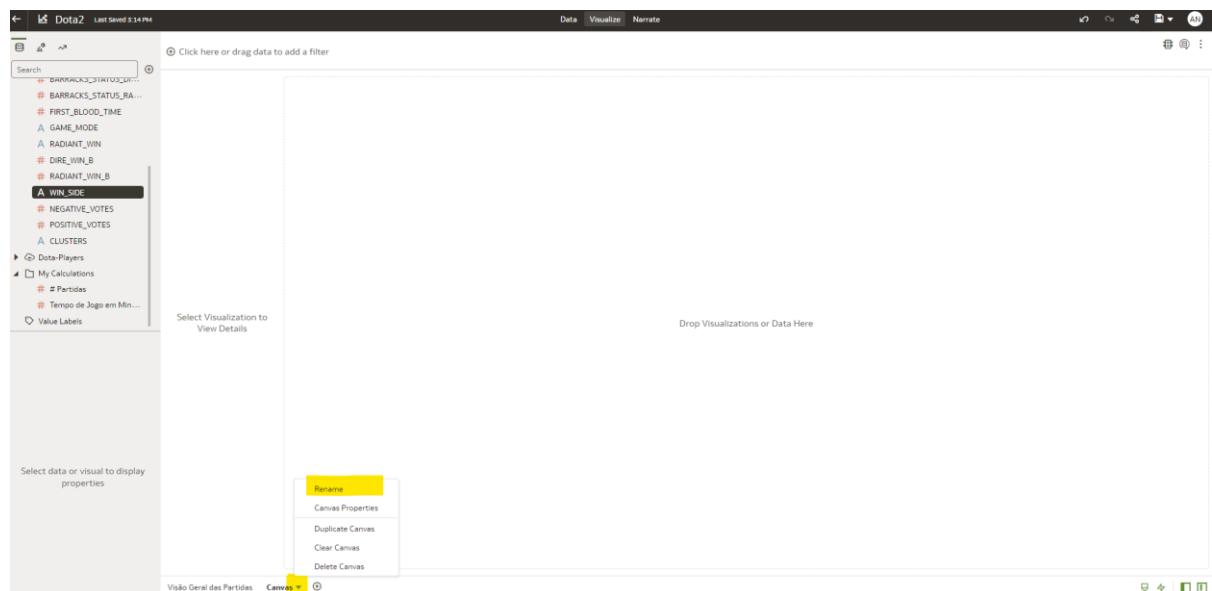
Insights:

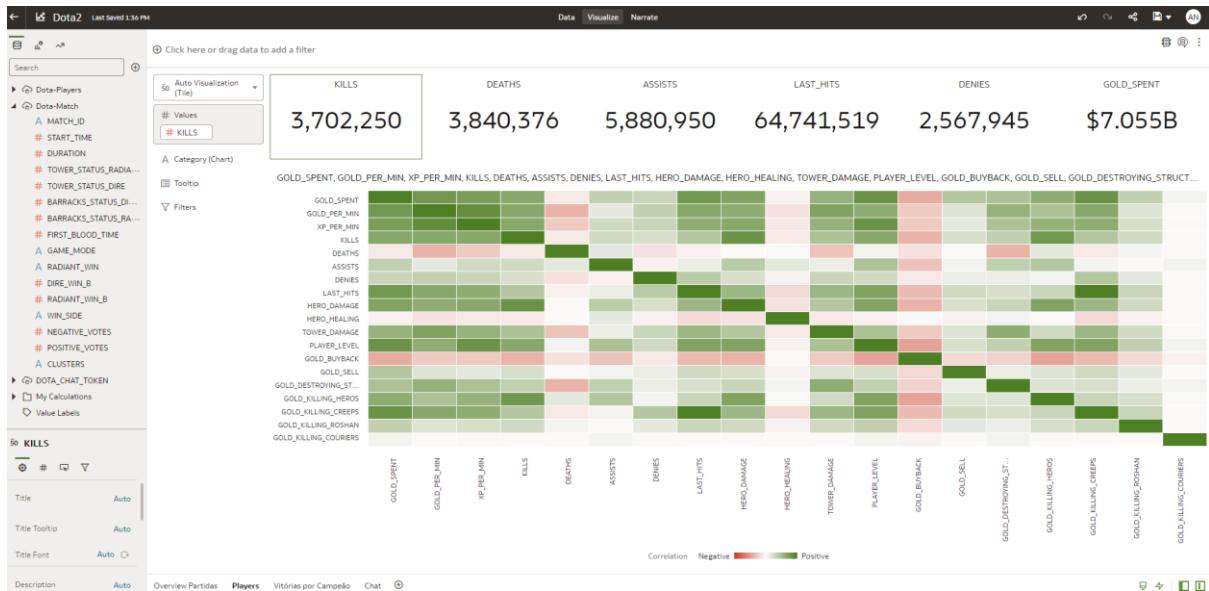
- Nesta análise já é possível observar que no mapa do Dota2 o Lado dos Radian tem uma certa vantagem não importando o modo de jogo e nem qual Cluster regional está sendo analisado
- Também é possível observar uma média de 41 minutos por partida e que as First Bloods acontecem bem rapidamente com uma média de 93 segundos dentre as 50 mil partidas analisadas.

Performance dos Jogadores

Este dashboard foi criado com a ideia de mostrar como a performance dos jogadores influenciam em uma vitória ou uma derrota, será observado como ganhar ouro, ganhar experiência, dar o último hit em uma creep ou “denar” uma creep pode influenciar positivamente ou negativamente no resultado da partida e também quais os números totais de toda a base de dados para se ter uma ideia de grandeza.

Primeiro para renomear os Dashboards/Canvas que compõem nosso projeto, o passo a passo é conforme abaixo.





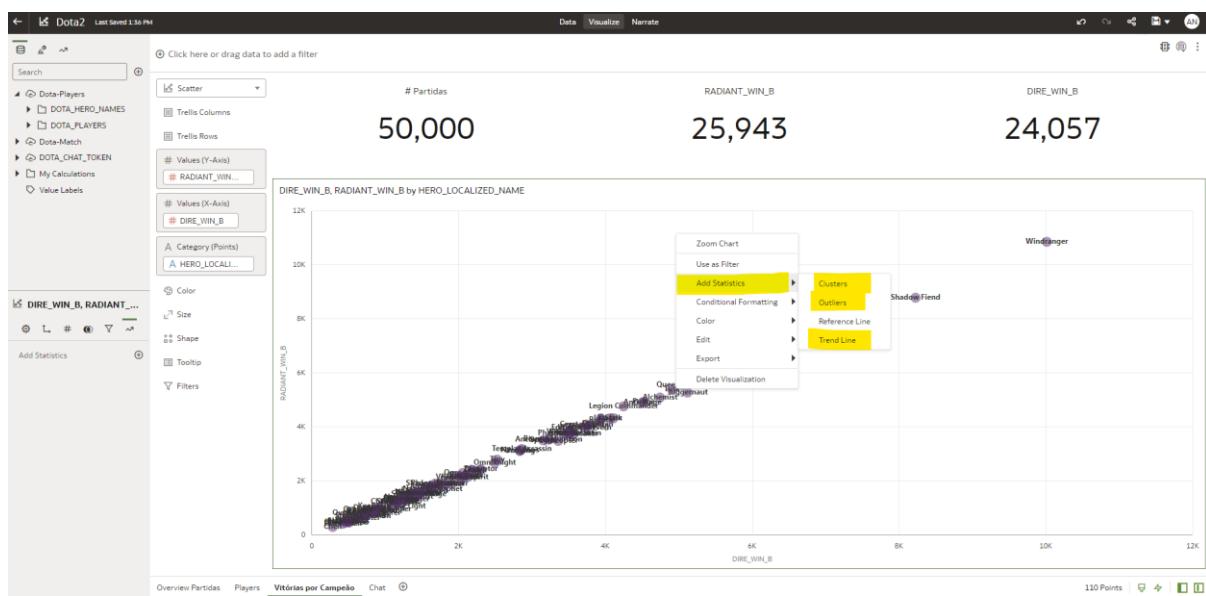
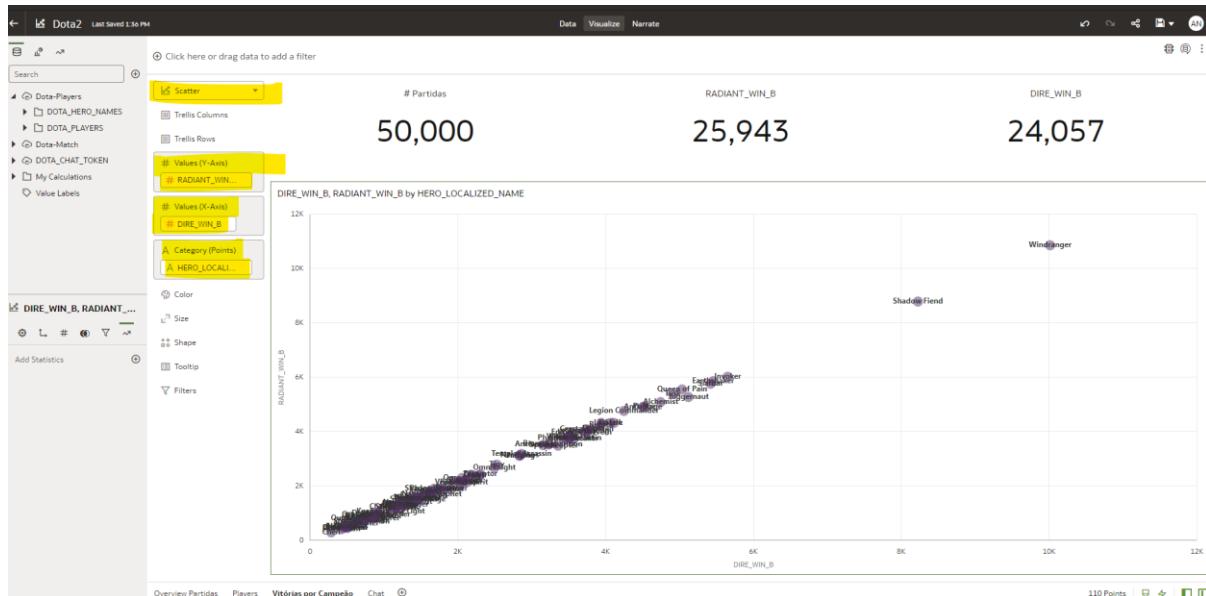
Insights:

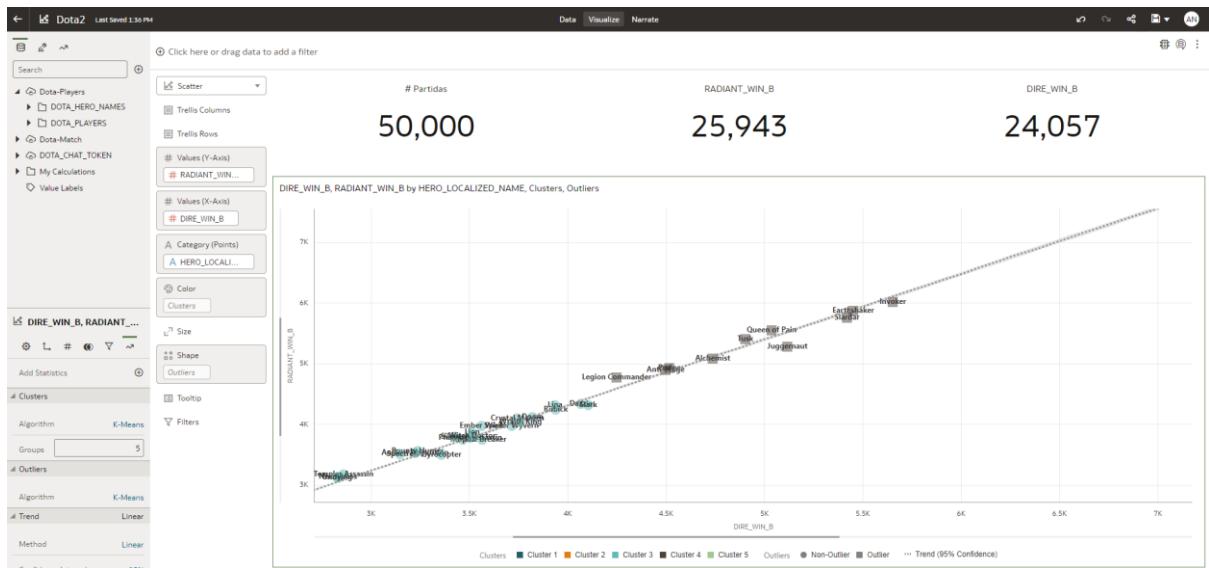
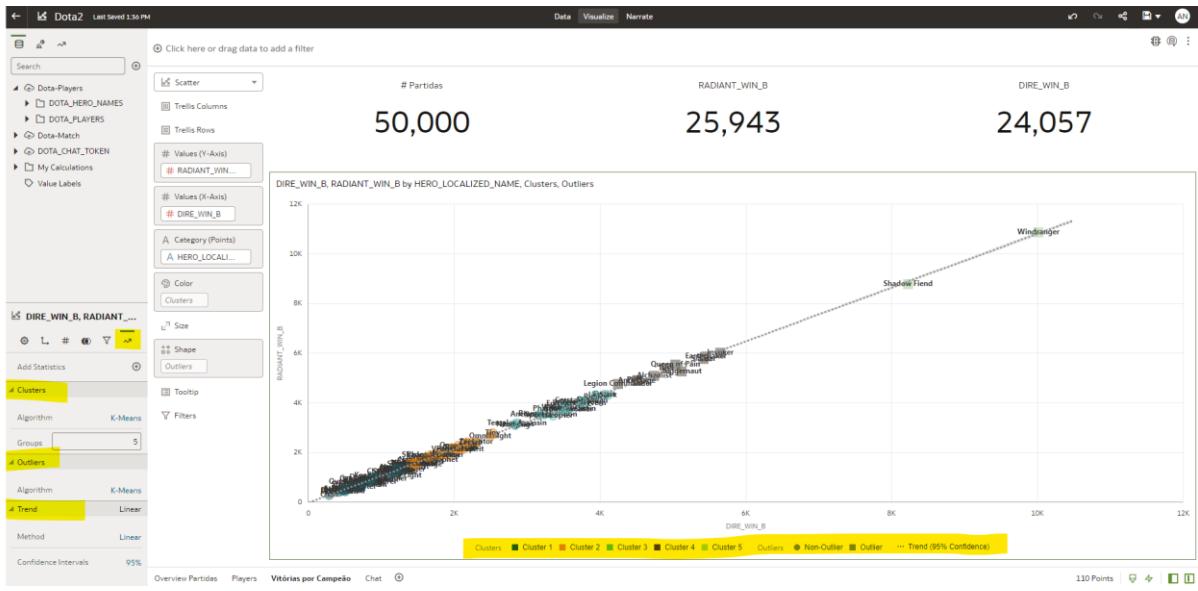
- Muitas perguntas podem ser respondidas com a Matriz de Correlação acima, uma coisa interessante é que quando o jogador morre no Dota2 ele perde um pouco de ouro que ele tem com ele, logo o dado de DEATH tem uma correlação negativa com o Ouro Gasto.
- É possível observar pelos dados que fazer o “Buyback”, ato de comprar o tempo no qual o herói está morto para já nascer imediatamente após alguma morte, não é tão bom assim para todas as métricas correlacionadas, logo uma dica seria usar essa possibilidade somente em último caso
- O ato de Denied é algo também exclusivo do dota, você nega o last hit na sua creep para que o inimigo não ganhe ouro e experiencia, isso melhora a sua performance convertendo em possibilidades de Kills para o seu time.

Há vários outros insights possíveis neste dashboard, será listado somente alguns.

Vitórias por Herói

Neste dashboard será observado a quantidade de vitórias por herói dependendo do lado que ele estava sendo Radiant ou Dire, a ideia também é clusterizar esses heróis, procurar outliers e criar uma reta de tendência para observar quais campeões aparecem mais do lado do Radiant ou do Dire.



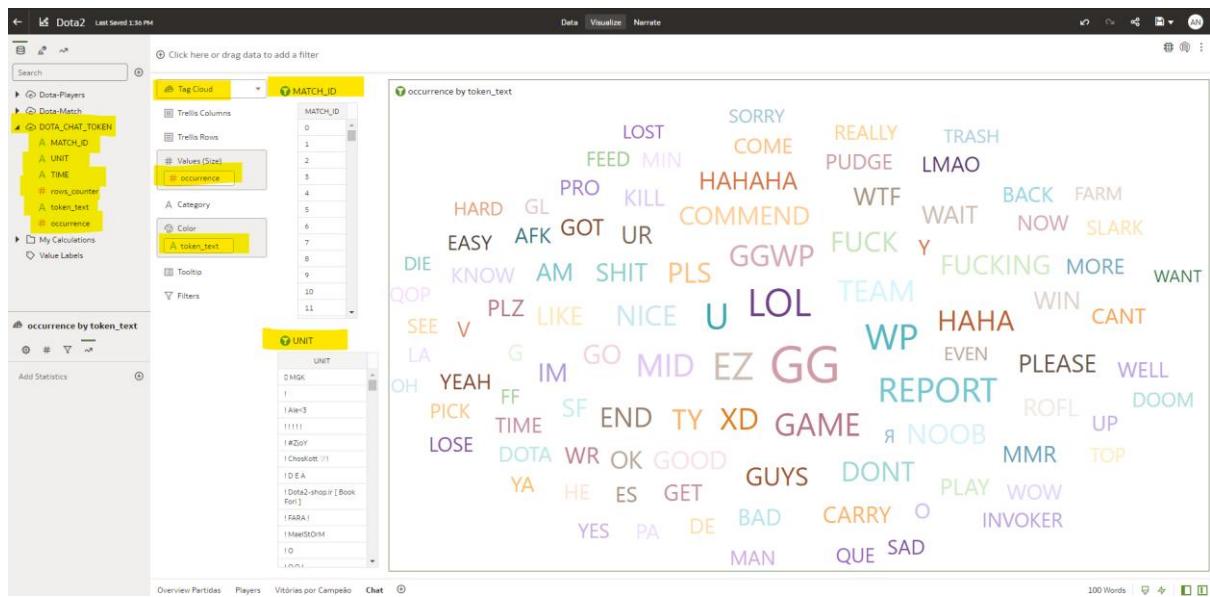


Insights:

- Há dois heróis que estão bem destacados de aparecer nas partidas, como a Windranger e o Shadow Fiend, durante o período de tempo que esta base de dados.
- Se for feito um Zoom no gráfico no Cluster 4, que são heróis que aparecem bastante mas não tanto quanto os do cluster 5 vemos que o Juggernaut ele ganha mais partidas pelo lado do Dire do que pelo lado Radiant, e o oposto ocorre com a Queen of Pain por exemplo.

Análise do Chat

Neste Dashboard será analisado quais são as palavras que mais surgem no chat do jogo durante uma partida. Aviso! A comunidade do Dota2 não é tão amistosa assim.



Insights:

- É possível observar que a comunidade do Dota é um pouco tóxica principalmente pelo surgimento de palavras como EZ (Easy), Report (Pedindo para reportar um jogador) e vários xingamentos.

Sinta-se livre para fazer suas explorações de dados, tirar os seus insights e aprender com dados.