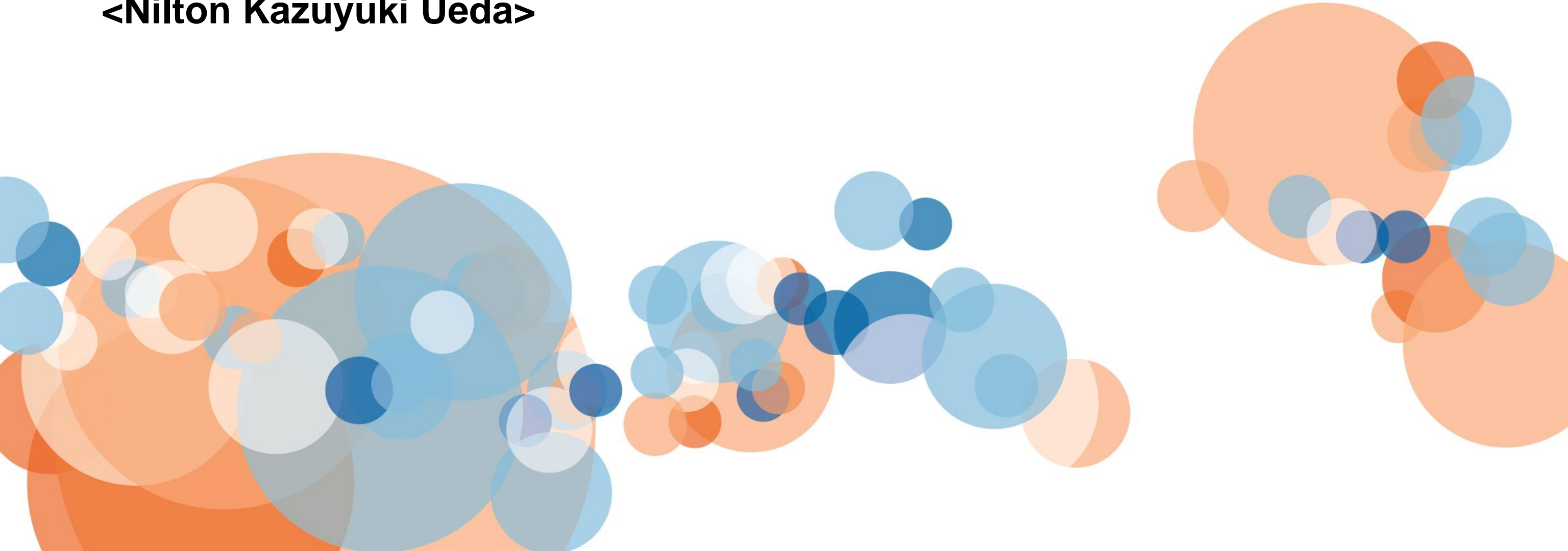




<Nilton Kazuyuki Ueda>



WHO I AM?



<Nilton Kazuyuki Ueda>

Contato: +55 11 96330-7703

Mail: kazu.ueda@hotmail.com


Portfólio: <http://tabsoft.co/2G5vGRy>

LinkedIn: <https://www.linkedin.com/in/niltonkazuyukiueda/>

Analista de BI & Analytics Sênior

Especialista em  **alteryx** | The Thrill of Solving

SCHEDULING

- **Business Intelligence Introduction**
 - What is Business Intelligence
 - Corporate Business Intelligence Architecture
 - **Tableau**
 - What is NOT Tableau?
 - What IS Tableau?
 - Gartner Magic Quadrant for Analytics Tools
 - Data Sources Connectors
 - How is Tableau?
 - How to scale Tableau in your company?
 - How much is Tableau?
 - Business Case
 - Demonstration Session
 - **TabPy**
 - Introduction
 - Tableau + Python
 - Installation
 - Starting the Service
 - Connecting the Service
 - Developing the service
 - Samples
 - Possible Errors
 - **Q&A**
- 





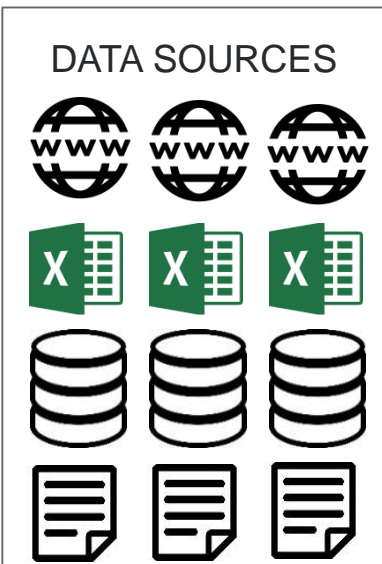
CORPORATE BUSINESS ANALYTICS ARCHITECTURE

COLLECTION

DATA ENGINEERING

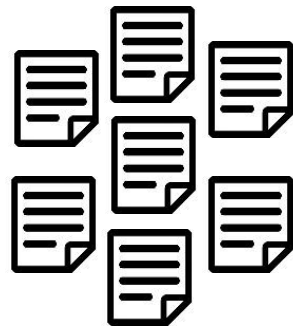
VISUALIZE and SHARE

END USER



SHELL

STAGING
(AS IS)



DATA CLEANSING
AND INTEGRATION



DATA
WAREHOUSE

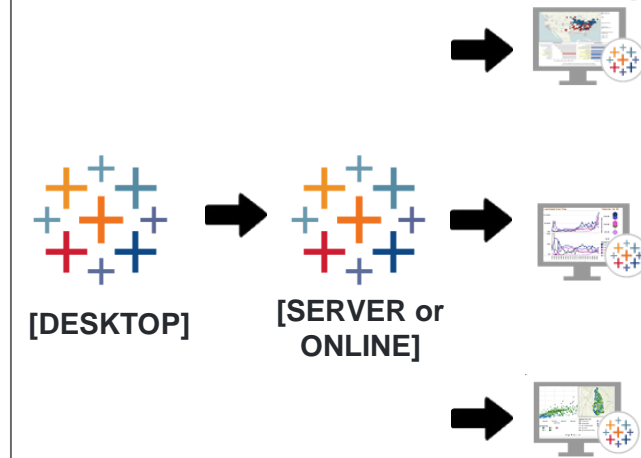


Live
connection or
schedule

ANALYTICS

[DESKTOP]

[SERVER or
ONLINE]



PIPELINE

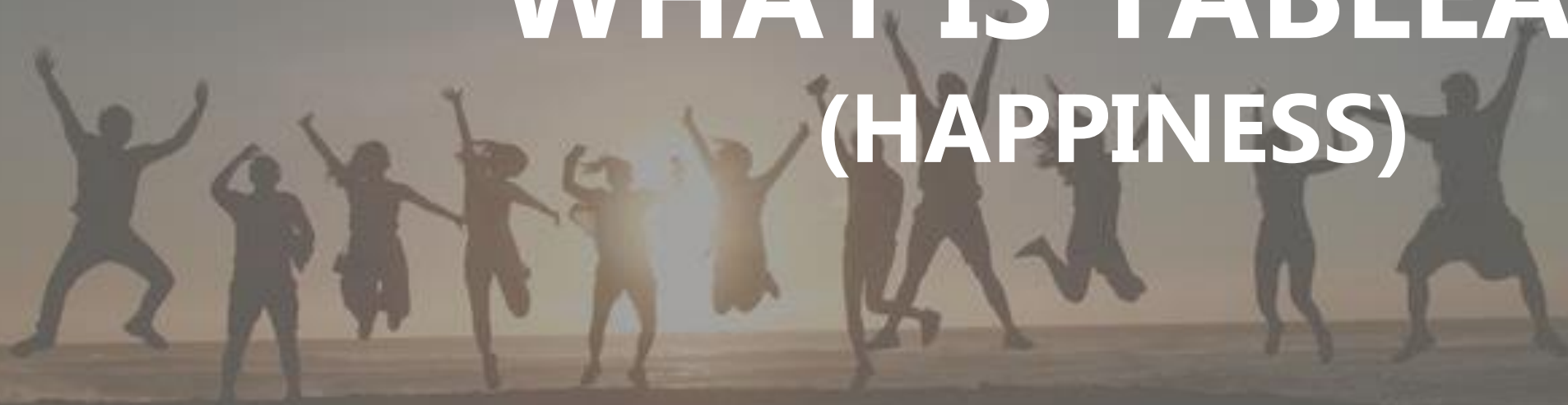
A cartoon illustration of Homer Simpson from 'The Simpsons'. He is depicted from the chest up, wearing his signature white short-sleeved shirt. His eyes are wide open, and he has a joyful expression. His two arms are raised high, with his fists clenched in a cheering gesture. The background is a solid, medium blue color. Overlaid on the center of the image is the text 'WE NEED DATA AND ANALYSIS' in a bold, white, sans-serif font.

WE NEED DATA AND ANALYSIS

WHAT IS **NOT** TABLEAU?

(CODE IS NECESSARY)

WHAT IS TABLEAU? (HAPPINESS)





WHAT IS TABLEAU?





WHY TABLEAU?

- Self Service** ➔
 - A TI não pôde responder com a rapidez necessária devido a prioridades conflitantes.
 - Os usuários precisavam da capacidade de fazer análises ad-hoc de maneira ágil.
 - Permite que os usuários explorem dados rapidamente e respondam a perguntas de negócios
- Intuitive** ➔
 - O Tableau ofereceu um conjunto de ferramentas que os usuários finais poderiam aprender rapidamente.
 - O Tableau oferece uma variedade de tutoriais on-line gratuitos / webinars que se mostraram muito eficazes para aumentar os usuários finais.
 - A ferramenta Tableau é estruturada de modo que as Melhores Práticas de BI sejam criadas “na caixa”.
- Rapid Development Cycle** ➔
 - As visualizações do Tableau podem ser desenvolvidas e editadas rapidamente
 - Capacidade de combinar diferentes fontes de dados sem intervenção de TI
 - Os usuários podem publicar conteúdo para seus públicos sem a necessidade de envolver a TI
 - Capacidade de desenvolver uma vez e implantar por meio de vários canais
- Security** ➔
 - O Tableau respeita a segurança de banco de dados existente e permite que a TI aproveite a infraestrutura de segurança existente
 - Elimina o risco de enviar arquivos por e-mail e possivelmente expor dados confidenciais

A person stands on a dark rock in the foreground, looking up at a vast night sky filled with stars. The Milky Way galaxy is visible as a bright, colorful band of light stretching across the sky, with hues of purple, pink, and orange. The person's silhouette is small against the immense scale of the universe.

Gartner®

Gartner, Inc. is an American research and advisory firm providing information technology related insight for IT and other business leaders located across the world.



GARTNER'S MAGIC QUADRANT FOR ANALYTICS (2018)

LEADERS and VISIONARIES



2015



2016



2017



2018



Conectar

A um arquivo

Microsoft Excel
Arquivo de texto
Arquivo JSON
Microsoft Access
Arquivo PDF
Arquivo espacial
Arquivo estatístico

Mais...

A um servidor

Tableau Server
Microsoft SQL Server
MySQL
Oracle
Amazon Redshift

Mais...

Fontes de dados salvas

Exemplo - Superloja
Indicadores Mundiais

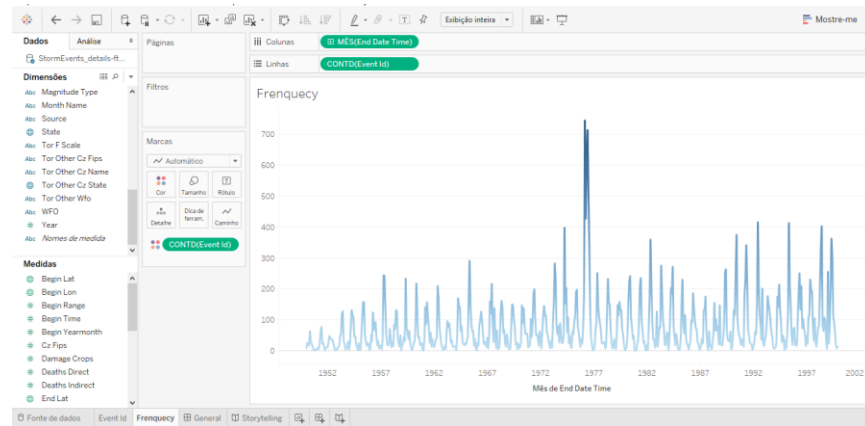
Pesquisar

Tableau Server	Hortonworks Hadoop Hive	Pivotal Greenplum Database
Action Matrix	HP Vertica	PostgreSQL
Action Vector	IBM BigInsights	Presto
Amazon Athena	IBM DB2	Progress OpenEdge
Amazon Aurora	IBM PDA (Netezza)	QuickBooks Online
Amazon EMR	Kognitio	Salesforce
Amazon Redshift	MapR Hadoop Hive	SAP HANA
Anaplan	Marketo	SAP NetWeaver Business Warehouse
Apache Drill	MarkLogic	SAP Sybase ASE
Arquitetura de dados	Microsoft SQL Server	SAP Sybase IQ
Box	Microsoft Analysis Services	ServiceNow ITSM
Cisco Information Server	Microsoft PowerPivot	Listas do SharePoint
Cloudera Hadoop	Microsoft SQL Server	Snowflake
Denodo	MonetDB	Spark SQL
Dropbox	Conector do MongoDB BI	Splunk
EXASOL	MySQL	Teradata
Firebird	OData	Teradata OLAP Connector
Google Analytics	OneDrive	Conector de dados da web
Google BigQuery	Oracle	
Google Cloud SQL	Oracle Eloqua	Outros bancos de dados (ODBC)
Planilhas Google	Oracle Essbase	

MORE THAN 80 DIFFERENT DATA CONNECTORS

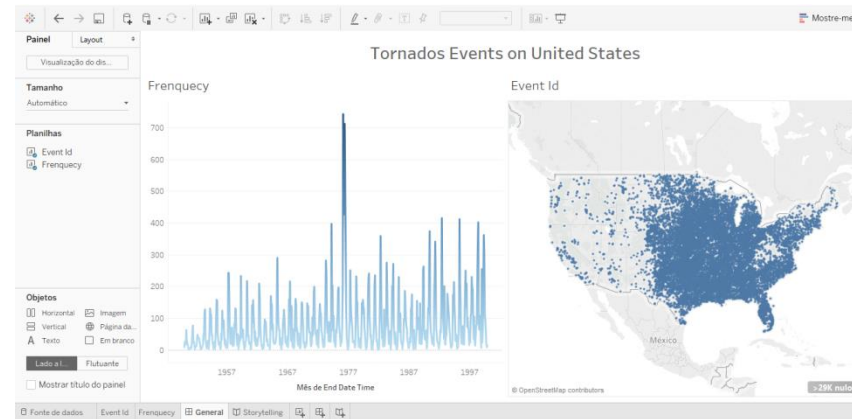


HOW IS TABLEAU?



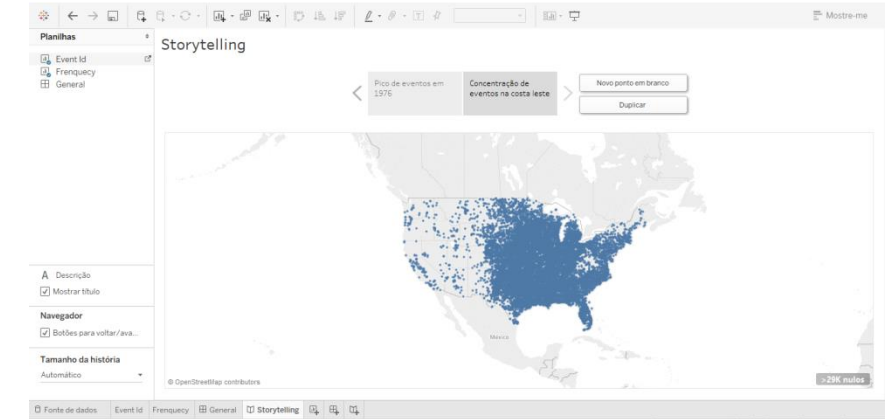
Worksheets

Where you create the graphics individually



Dashboards

Where you consolidate the graphs and create integrated analyzes



Storytelling

Where do you create your executive presentations and storytelling





HOW TO SCALE TABLEAU IN YOUR COMPANY?



Tableau Desktop

Crie visualizações de dados impactantes usando quaisquer tipos de dados. Rápido e fácil.

- Visualizações de dados e painéis interativos.
- Compatível com centenas de fontes de dados.



Tableau Server

Compartilhe painéis e dados. Colabore com governança.

- Mantenha os dados protegidos em seus servidores.
- Use o Tableau Mobile para acessar dados quando você estiver em trânsito.



Tableau Online

Elimine a etapa de configuração de hardware. O Tableau Online é uma versão hospedada do Tableau Server.

- Confiável, seguro e sempre atualizado.
- Use o Tableau Mobile para acessar dados quando você estiver em trânsito.





HOW MUCH IS TABLEAU?

CREATOR

USD \$70 ^{R\$246,40}

Monthly Billing

Yearly USD \$840 ^{R\$2.956,80}

Discover insights with a powerful suite of products that support your end-to-end analytics workflow.

Includes:

Tableau Desktop
Tableau Prep
Tableau Server

EXPLORER

USD \$35 ^{R\$123,20}

Monthly Billing

Yearly USD \$420 ^{R\$1.478,40}

Explore trusted data and answer your own questions faster with full self-service analytics.

Includes:

Tableau Server

VIEWER*

USD \$12 ^{R\$42,24}

Monthly Billing

Yearly USD \$144 ^{R\$506,88}

View and interact with dashboards and visualizations in a secure, easy-to-use platform.

Includes:

Tableau Server

(*) Minimum 100 users

A person's arm is visible on the left side of the frame, resting on a desk. In the background, a laptop is open, displaying a document. To the right of the laptop is a rolled-up document or notebook. The entire scene is overlaid with a semi-transparent dark filter.

www.tableau.com/academic
(only students and teachers)



BUSINESS CASE

AVIATION

DRUGSTORE

TELECOM

INSURANCE

RETAIL



<https://goo.gl/JmSKwp>

<https://goo.gl/nqtEHo>

<https://goo.gl/KssUyR>

<https://goo.gl/wex9pR>

<https://goo.gl/ZsDYkr>



DEMO SESSION

LET'S SEE IN PRACTICE?



TABLEAU + PYTHON = TABPY



+ tableau®

+



python™



INTRODUÇÃO

Essa nova funcionalidade do Tableau 10.2 é uma conexão remota ao servidor Python.

O Tableau Python Server (TabPy) é um framework que possibilita a conexão e execução remota de códigos Python Via Tableau.

Este framework possui um servidor construído em Tornado e via REST API conseguimos utilizar todo o poder e flexibilidade do Python no Tableau.

Neste servidor já temos instaladas várias bibliotecas e módulos para pronta utilização mas conseguimos instalar mais módulos e bibliotecas caso necessário.

Ao conectar no Tabpy o tableau consegue enviar códigos para execução e utilizar este retorno nas visualizações criadas.



TABLEAU 10.2 VS PYTHON (TABPY)

Qual ganho de utilizar essa nova funcionalidade?



Análises sentimental de texto;



Modelos preditivos mais complexos;



Todo o poder e versatilidade da linguagem;



O pacote vem com diversos pacotes por padrão e é possível adicionar de acordo com a necessidade;



É possível executar de forma simples e integrada modelos personalizados de Machine Learning;



INSTALAÇÃO

Para instalação, temos disponível um script para instalação.

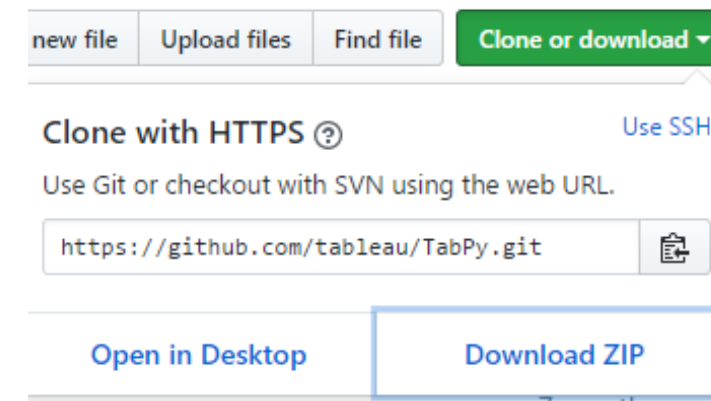


TabPy-master.zip

Dentro deste arquivo temos dois setups, setup.sh, utilizada para ambientes Linux e setup.bat, utilizado para ambientes Windows. Encontre qual é o arquivo que você necessita e execute-o.

O que esse script realiza:

- Este script baixa e instala o Anaconda (Uma das plataformas Python)
- Cria um ambiente Python com o nome de Tableau-Python-Server
- Realiza a instalação de módulos e bibliotecas prontos para utilização



Obs.: Caso não consiga abrir o arquivo anexo, entrar no link <https://github.com/tableau/TabPy> e clicar em: clone or download >> Download ZIP



INSTALAÇÃO

Se a mensagem abaixo aparecer, isso indica a correta instalação do Servidor e qual a porta você precisa utilizar para realizar a conexão:

“Done initializing tabpy
Web service listening on port **9004**”

Observações e cuidados:

- A conexão ao Tabpy funciona melhor com a versão 2.7 Python, caso haja outra versão instalada, é necessário instalar esta também.
- Caso esteja instalada em sua máquina local, o serviço precisa ser iniciado toda vez que for utilizar a conexão.



INICIANDO O SERVIÇO

Para conseguir se conectar ao TabPy o serviço precisa estar iniciado. Caso não esteja, basta entrar no diretório abaixo:

C:\Users\USER\Anaconda\envs\Tableau-Python-Server\Lib\site-packages\tabpy_server\ - Alterar USER para seu user windows

Executar o arquivo: startup.bat.

Este script irá iniciar o serviço e informar qual porta está aberta para conexão.

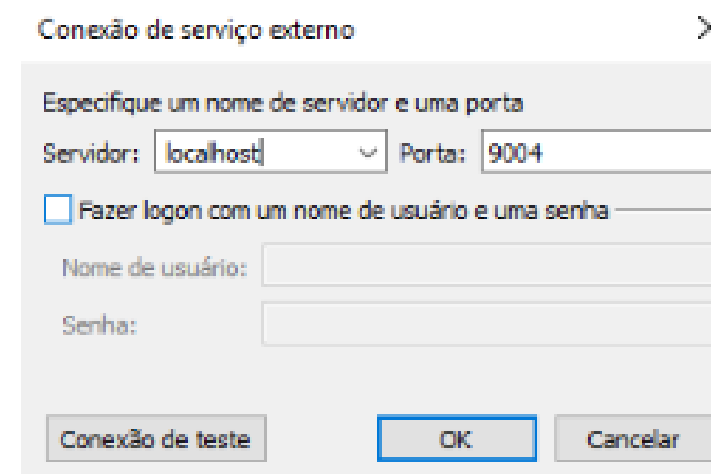
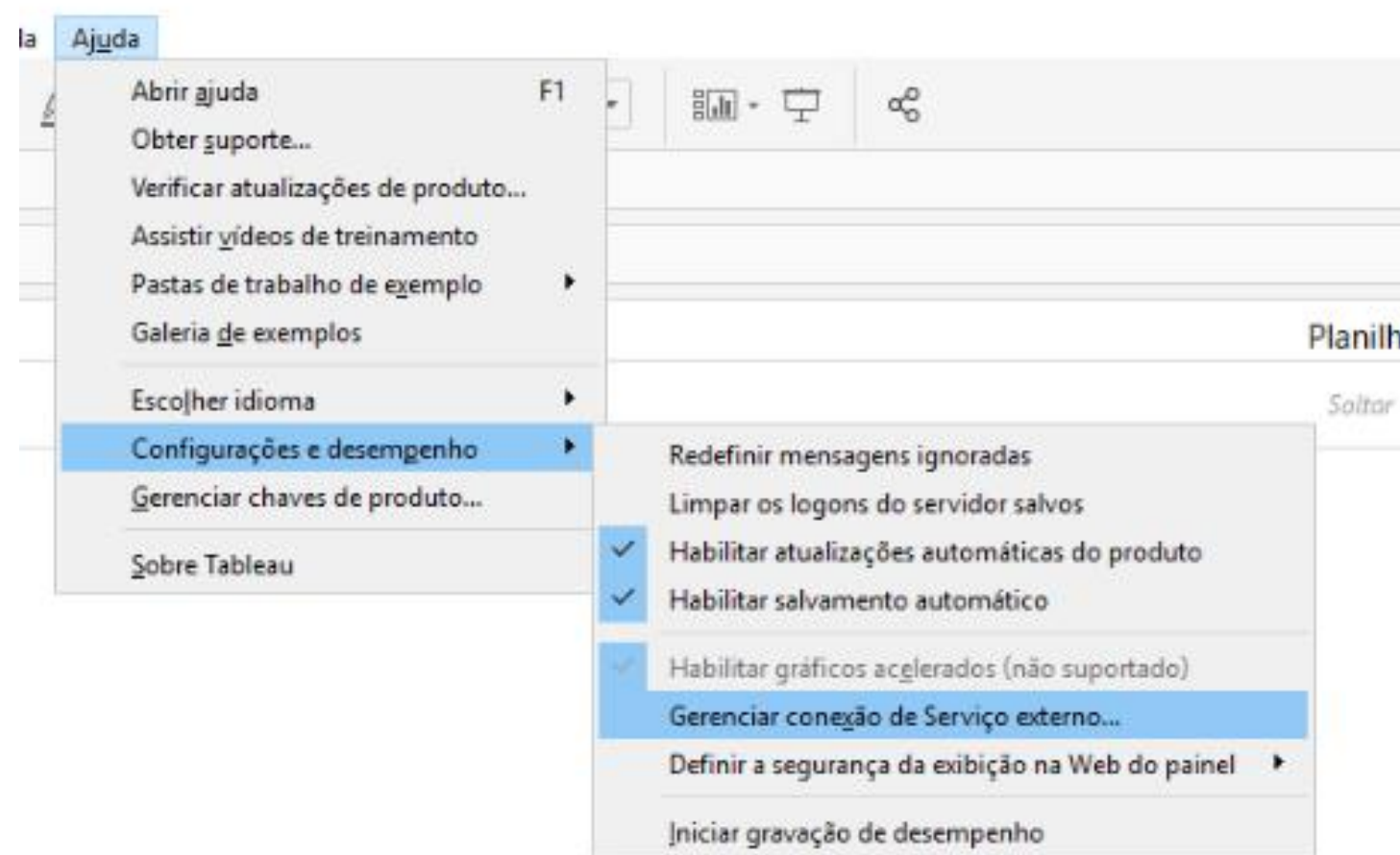
```
C:\> C:\WINDOWS\system32\cmd.exe

Initializing TabPy...
Done initializing TabPy.
Web service listening on port 9004
```



CONECTANDO

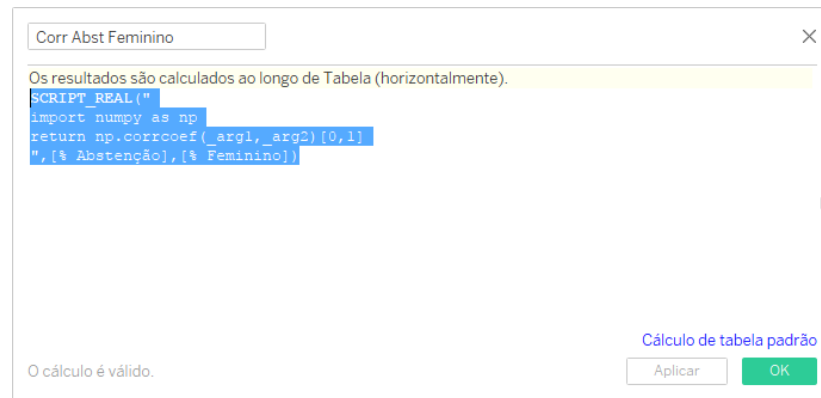
Clicar em Ajuda > Configurações e desempenho > Gerenciar conexão de Serviço Externo e adicionar “localhost” no Servidor e o número da porta que apareceu ao final da instalação





DESENVOLVENDO

Para executar códigos no Python via tableau temos que usar um script que passa o código parametrizado a ser executado e os parâmetros para execução. Ex.:



Para cada tipo de dado que é retornado do Python temos um script para recepção:

SCRIPT_BOOL(): Tipo de dado no retorno *booleano* (*T ou F; 1 ou 0*)

SCRIPT_INT(): Tipo de dado no retorno inteiro (42;1123581321;)

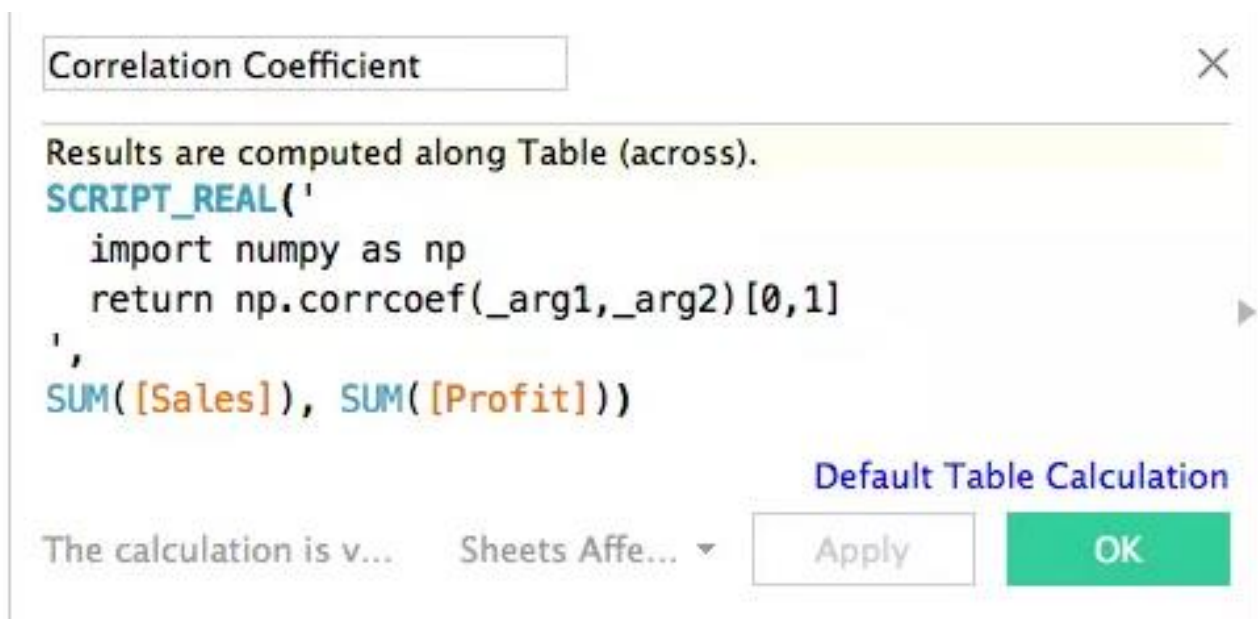
SCRIPT_REAL(): Tipo de dado no retorno float (3,14 ; 0,37373737)

SCRIPT_STR(): Tipo de dado no retorno cadeia de caracteres (#TUG SP)



EXEMPLO

Utilizando a base da superstore, conseguimos criar o campo abaixo:



Essa função de correlação (`np.corrcoef`) retorna valores entre -1 até 1, indicando a relação de proporcionalidade entre duas métricas. Ou seja, quanto mais próximo de 1, mais diretamente proporcional são as variáveis e quanto mais próximo de -1, mais indiretamente proporcional são as variáveis.



EXEMPLO

Crie uma visão de dispersão com **Vendas**, **Lucro**, separadas entre **Sub-Categoria** e **Região**.

Em Marcas, Detalhe, arraste o campo de **Customer ID** ou **Customer Name**.

Arraste para a cor o campo **Correlation Coefficient**.

Neste momento temos o cálculo de correlação de vendas e lucro por cliente dentro de cada região e subcategoria de produto. Quanto menor o “grão” no detalhe da visão, melhor, pela quantidade de informações para comparação.

Columns	Sub-Category	SUM(Profit)
Rows	Region	SUM(Sales)

Sheet 21

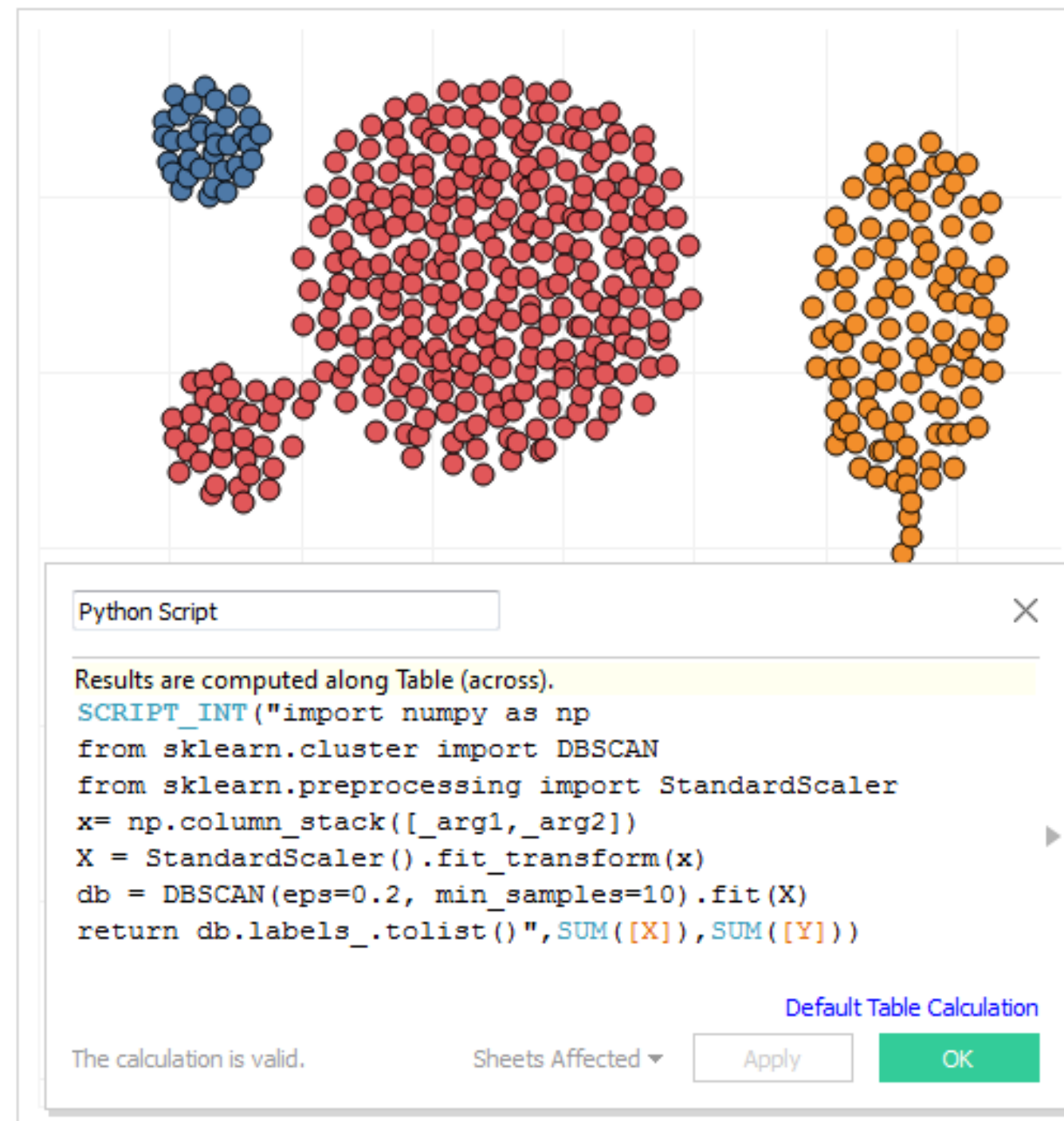




EXEMPLO

Em um outro exemplo podemos usar a library **numpy** combinada com a library **sklearn** para realizar análises de dados considerando o algoritmo de cluster **k-means**.

Basta o serviço está disponível e você deixar a conexão ativa para usar todo o poder que o Python oferece.





POSSÍVEIS ERROS

Possíveis erros que podemos encontrar:

- O TabPy Funciona melhor com a versão 2.7 do Python. Caso tenha alguma versão instalada na sua máquina, isto pode gerar um conflito. Caso isso ocorra, Desinstale qualquer versão existente e execute novamente o instalador.
- Todo cálculo criado para execução no Python deve mandar uma quantidade de registros e retornar esta mesma quantidade de registros. Logo, o cálculo de tabela deve ser feito baseado no menor nível presente na visão. Em nosso exemplo, caso seja apresentado algum erro no momento em que for adicionado o campo de correlação na visão, clique com o direito no campo, selecione **Editar cálculo de tabela...** e selecione as opções conforme a figura ao lado.

Cálculo de tabela
Correlation Coefficient

☐ Nulo será gerado se não houver valores suficientes

Calcular usando

- Tabela (horizontalmente)
- Tabela (abaixo)
- Tabela (horizontalmente e depois abaixo)
- Tabela (abaixo e depois horizontalmente)
- Painel (horizontalmente)
- Painel (abaixo)
- Painel (horizontalmente e depois abaixo)
- Painel (abaixo e depois horizontalmente)
- Célula
- Dimensões específicas**

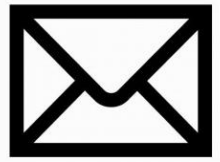
☒ Customer ID
☐ Region
☐ Sub-Category



REFERÊNCIAS

- <https://github.com/tableau/TabPy>
- <https://www.tableau.com/about/blog/2017/1/building-advanced-analytics-applications-tabpy-64916>
- <https://www.tableau.com/about/blog/2016/11/leverage-power-python-tableau-tabpy-62077>
- <https://github.com/tableau/TabPy/blob/master/server.md>

Thanks for watching!



kazu.ueda@hotmail.com



+55 11 96330-7703



<https://www.linkedin.com/in/niltonkazuyukiueda/>



<https://public.tableau.com/profile/nilton.kazuyuki.ueda/>