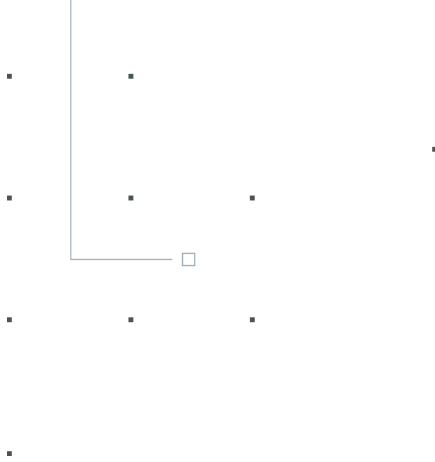
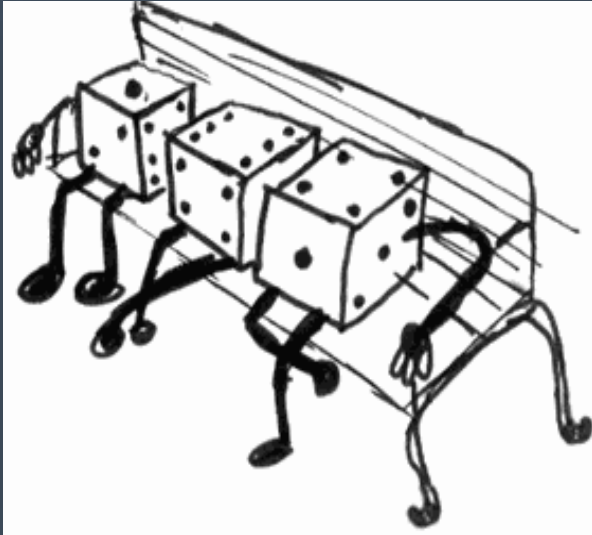


FIAP

NBA



# Data Ecosystems



## APRESENTAÇÃO

- Apresentação da professora;
- Objetivo da Disciplina;
- Conteúdo Programático;
- Metodologia Aplicada;
- Avaliação;
- Referências Bibliográficas.



## Tassiana Rugoni de Campos

*Consultora em Business Intelligence e Big Data*

Especialista em Sistemas de Informação com ênfase em Tecnologia GIS, Banco de Dados e Marketing Direto.

Coordenadora do MBAs: Arquitetura de Soluções / Gestão de Tecnologia da Informação – Fiap On.

Professora nos MBAs de ASO, BI, Big Data, IA, Eng. Software, Eng. De Dados. Professora de Redes de Computadores e Programação Avançada para turmas de graduação de engenharia.

Professora responsável por curso de graduação EAD em ADS.

Autora de livros didáticos sobre:

- Qualidade de Software – CMMI;
- Desenvolvimentos: Mobile Android (Java), Python, Web (PHP/SQL) e Jogos (Unity – C#);
- Arquitetura de Data Lake (Hadoop) e Streaming de Dados (Flume, Sqoop, Kafka, Nifi).

Experiência profissional na gestão de equipe de infraestrutura.

E gestão de equipe de projetos para soluções de Banco de Dados, BI e DW.

### Contatos:

[proftassiana.campos@fiap.com.br](mailto:proftassiana.campos@fiap.com.br)

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

<http://lattes.cnpq.br/4319895020677966>

- [illegible]

# RESUMO DO CONTEÚDO PROGRAMÁTICO

- Contextualização do cenário atual de dados (compreensão do mercado);
- Conceitos sobre BI (DW, Data Marts, Data Mining, Modelagem dimensional);
- Conceitos sobre Big Data (Data Lake, Data LakeHouse, Hadoop, Map Reduce, HDFS);
- Bancos SQL versus NoSQL;
- BI versus Big Data;
- Ingestão e Streaming de Dados;
- Engenharia e Desenvolvimento de Soluções Analíticas (Sensing, Collection, Wrangling, Analysis, Storage);
- Abordagens Arquiteturais (Lambda, Kappa, Data Mesh);
- Arquitetura Física, Distribuída, Em Nuvem.
- Plataformas em Cloud: Databricks, Snowflake, Trifacta, Data Robot.

## METODOLOGIA APLICADA

- Aula expositiva para apresentação do conteúdo através da exposição de slides.
- Aula expositiva para apresentação do conteúdo através da exposição de slides.
- Debates em sala de aula visando troca de experiências.
- Aplicação de estudo de caso para desenvolvimento de uma solução de Trade Marketing.
- Hands On para solução do Case.



# AVALIAÇÃO

- Projeto: Case de Trade Marketing:





# AVALIAÇÃO

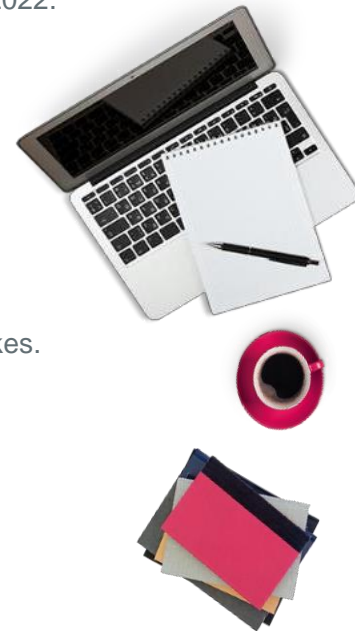
- Teste de Conhecimento:



Teste seus conhecimentos!  
*Clique aqui e responda.*

# REFERÊNCIAS BIBLIOGRÁFICAS

- STOBIERSKI, Tim. 5 KEY ELEMENTS OF A DATA ECOSYSTEM. Disponível em: <<https://online.hbs.edu/blog/post/data-ecosystem>>. Acesso em 20/09/2022.
- MACHADO, Felipe N. R. Big Data – O Futuro dos dados e aplicações. São Paulo: Erica, 2018.
- CASTRO, L., FERRARI, D. Introdução à mineração de dados. São Paulo: Saraiva, 2016.
- HAGSTROEM, M., ROGGENDORF, M. A smarter way to jump into data lakes. Disponível em: <<https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/a-smarter-way-to-jump-into-data-lakes>> . Acesso em 28/01/2019.
- INMON, William. Como usar o Data Warehouse. São Paulo: Makron Books, 1999.



## REFERÊNCIAS BIBLIOGRÁFICAS

- KIMBALL, R. The Data Warehouse Toolkit: guia completo para modelagem dimensional. Rio de Janeiro: Campus, 2002.
- BARBIERI, Carlos. BI2- Business Intelligence Modelagem & Qualidade. Rio de Janeiro, Elsevier, 2011.
- Machado, F. N. Tecnologia e projeto de Data Warehouse: uma visão multidimensional. São Paulo: Érica, 2004.
- TURBAN, E.; SHARDA, R.; ARONSON, J.; KING, D. BUSINESS INTELLIGENCE: um enfoque gerencial para a inteligência do negócio. Porto Alegre: Bookman, 2009.
- ELMASRI, R; NAVATHE, S.B. Sistemas de Banco de Dados. São Paulo: Pearson, 2011.



## REFERÊNCIAS BIBLIOGRÁFICAS

- LEE, J.; WEI, T; MUKHIYA, S. **Hands-On Big Data Modeling**. UK, Packt Publishing, 2018.
- HOBERMAN, S. **Data Modeling Made Simple with ER/Studio Data Architect: Adapting to Agile Data Modeling in a Big Data World**. EUA, Technics Publications, 2015.
- TILLMANN, G. **Usage-Driven Database Design: From Logical Data Modeling through Physical Schema Definition**. EUA, Apress, 2017.
- HILLS, T. **NoSQL and SQL Data Modeling: Bringing Together Data, Semantics, and Software**. EUA, Technics Publica, 2016.



- Contextualizar o Cenário de Dados;
- Entender o que é um ecossistema de dados;
- Mapear a camada de sensing;
- Conhecer o case de Trade Marketing;
- Desenhar o Business Process Matrix.

- Contextualizar o Cenário de Dados;
- Entender o que é um ecossistema de dados;
- Mapear a camada de sensing;
- Conhecer o case de Trade Marketing;
- Desenhar o Business Process Matrix.

# Uma Cultura Data Driven



Como a evolução tecnológica impacta no modo com o qual lidamos com os dados?



# Crescimento Exponencial da Captura de Dados

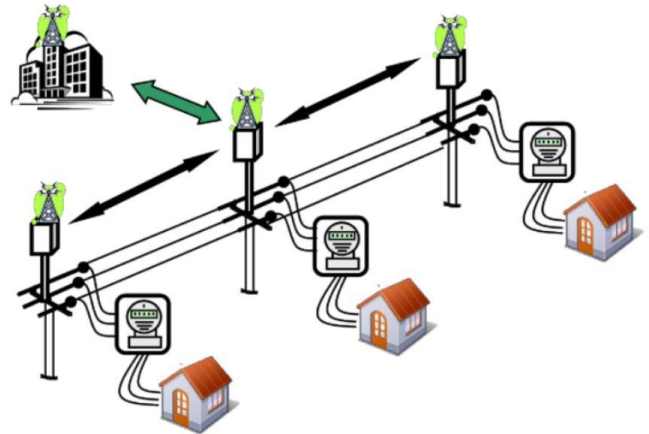
*Novas Fontes e Tipos de Dados*



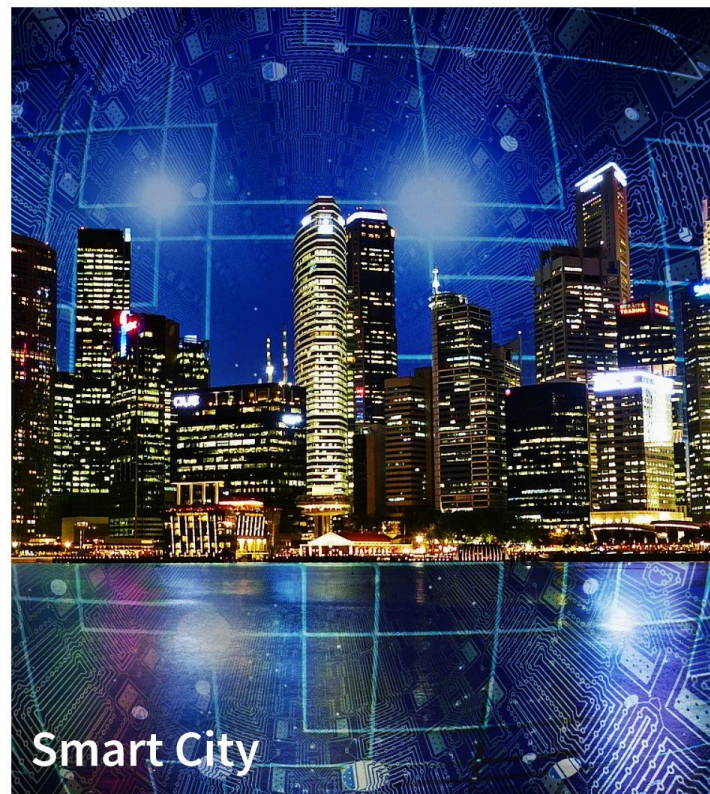
1 - Sensoriamento de Objetos e Captura de Eventos



## Smart Grid



Smart City





# Centrica implanta cluster Hortonworks em grande escala para impulsionar negócios

A empresa começou a trabalhar com a fabricante de caldeiras de aquecimento central Worcester Bosch para coletar diagnósticos de caldeiras remotamente. Com essa caldeira conectada, o sistema de termostato de internet Hive da empresa começa a se tornar um serviço, disse Hussain.

Como a Computer Weekly informou anteriormente, a divisão britânica de gás da empresa usou painéis do QlikView no data lake do Hadoop para suportar seu programa de medidores inteligentes.

O QlikView fornece análise de autoatendimento sobre um lago de dados do Hadoop. O data lake contém mais de nove bilhões de registros e coleta dados de fontes, incluindo mais de 150 tabelas SAP.



# Modernas arquiteturas de telecom integradas com Hadoop

Muitas das maiores empresas de telecomunicações do mundo usam o [Hortonworks Data Platform](#) (HDP) para gerenciar seus dados. Por meio da parceria com essas empresas, aprendemos como nossos clientes usam o HDP para melhorar a satisfação do cliente, fazer melhores investimentos em infraestrutura e desenvolver novos produtos.

A parceira da Hortonworks, [Teradata](#), recentemente deu alguns [exemplos de casos de uso neste vídeo](#) sobre como a Verizon Wireless usa o Teradata em combinação com a Hortonworks Data Platform para manter a queda de seus clientes abaixo de 1%.

Rob Smith, diretor executivo de TI da Verizon Wireless, descreve como sua equipe usa sua plataforma de descoberta para melhorar as interações com os clientes:

- Identificando maneiras melhores de se comunicar com os clientes sobre seus pagamentos
- [Analisando as mídias sociais para entender melhor os sentimentos dos clientes sobre as mudanças na política da Verizon](#)
- Adaptar as comunicações de marketing às necessidades individuais de cada cliente

# Análise de Sentimento no Twitter

[illegible]

・音ゲーマー見習い(たのしかったダンエボ 武寺:2554-8213)

[illegible]

## ANALISAR REGISTROS DE DETALHES DE CHAMADAS (CDRS)

Os Telcos executam perícia em chamadas perdidas e baixa qualidade de som, mas os registros de detalhes de chamadas fluem a uma taxa de milhões por segundo. Esse alto volume dificulta o reconhecimento de padrões e a análise de causa raiz, e muitas vezes eles **precisam acontecer em tempo real**, enquanto o cliente espera na linha ouvindo para manter a música. Atraso provoca atrito e prejudica as margens de serviço.

O **Apache Flume** **pode ingerir milhões de CDRs por segundo** no Hadoop, enquanto o **Apache Storm** processa esses dados em tempo real e identifica quaisquer padrões preocupantes.

O HDP facilita a retenção de dados a longo prazo para análise de causa raiz, mesmo anos após o primeiro problema. Essa análise de CDR pode ser usada para melhorar continuamente a qualidade da chamada, a satisfação do cliente e as margens de serviço.

**O que um cientista precisa saber para projetar soluções cuja análise e mineração de dados sejam vitais?**

EasyRetro

Perfil do Cientista

★ Prime Directive Tassiana Rugoni de Campos

Set the context of the board here...

Search Sort by order Add Share Settings

Soft Skills Hard Skills Percentual Negócio/Técnico

Comunicação Python 50/50

<https://easyretro.io/publicboard/RkZMDraAlgPWRSSaMmSYqmFZrHJ3/b3610eb2-8f1c-491c-8fb0-35d76522c3aa>



# Data Ecosystems

## 5 key elements

# COMPONENTES DO ECOSSISTEMA DE DADOS





## SENSING

- Dados Proprietários:



- Dados Legados:



- Dados Abertos:





**HANDS ON**



# CASE DE TRADE MARKETING

## Mobilidade gerando resultados

Hoje, muitas empresas possuem colaboradores externos que visitam clientes ou pontos de venda (PDV), e existe a dificuldade de medir a eficiência e o desempenho do seu trabalho. Mais do que simplesmente visualizar a execução operacional das atividades das equipes em campo, o importante é entregar a eles **informações estratégicas** que lhes permitam executar as ações nos PDVs com **maior eficiência**.



## Uma solução de Trade:



### Acompanha rotas e visitas geolocalizadas

Na rua, equipes de Trade Marketing podem manter seus gestores atualizados sobre o status de suas visitas em tempo real, garantindo mais eficiência e visibilidade de toda a equipe.



### Formulários customizados para coleta de dados

Transforme formulários complexos em papel em telas de simples preenchimento e intuitivas. Agilize o processo de coleta de dados e auditoria no PDV.



### Apresenta Dashboards e KPIs alimentados em tempo real

Determine quais indicadores precisam ser monitorados e deixe que nosso aplicativo colete os dados e os atualizem em tempo real. Visualize apenas o que é realmente importante de forma prática.



### Valida Estoque Virtual

Aumente suas vendas fazendo com que informações de estoque virtual cheguem à mão de seus promotores e sejam verificadas.



### Pesquisa Preços e Monitora a Concorrência

Monitore preços e informações da concorrência diariamente, visualizando os resultados desse trabalho em painéis adaptados às mais diversas necessidades.



### Captura e audita conquistas de espaços e materiais

Motivar equipes através de campanhas de incentivo que utilizam como base conquistas no PDV passou a ser uma tarefa simples com nosso aplicativo.

# INDICADORES E DIMENSÕES



•

□

+

+

•

•

•

•

•

□

•

•

•

•

# IDENTIFICANDO INDICADORES E DIMENSÕES

Indicadores



## O ATRIBUTO VARIA CONTINUAMENTE:

- Saldo
- Estoque
- Custo
- Vendas



## O ATRIBUTO é PERCEBIDO COMO CONSTANTE OU DISCRETO:

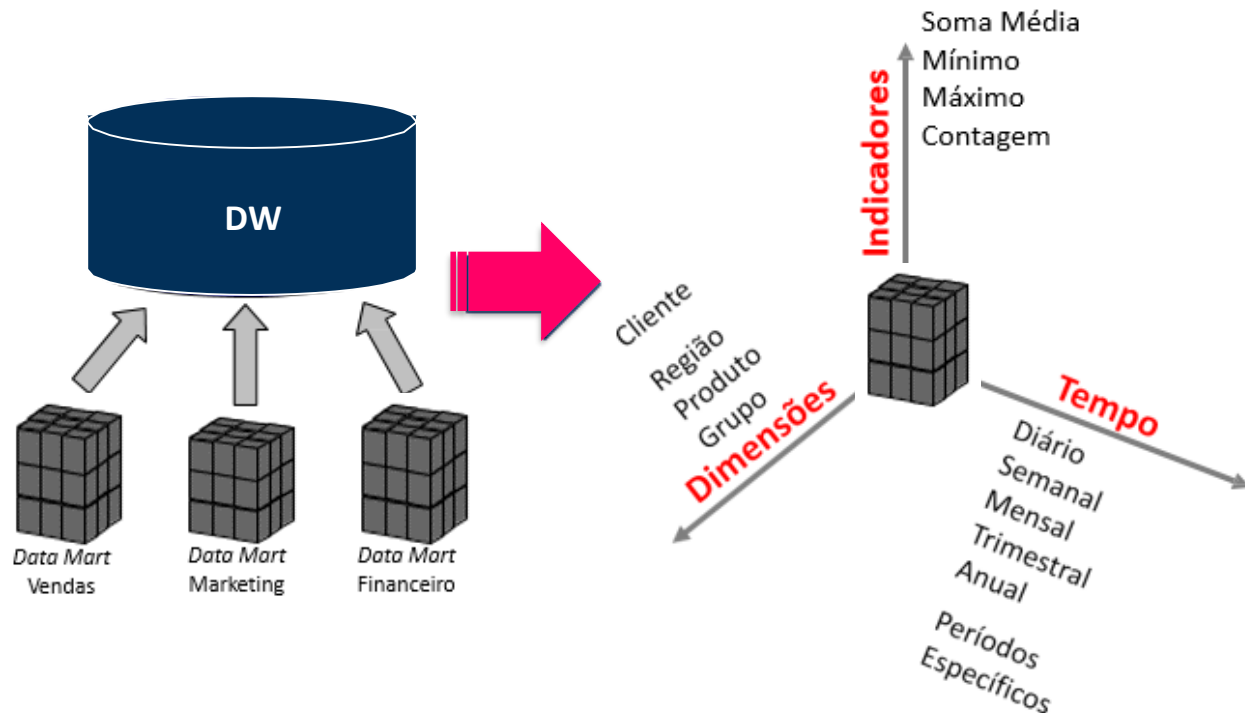
- Produto
- Localização
- Tempo (hora, dias, meses)



Dimensões

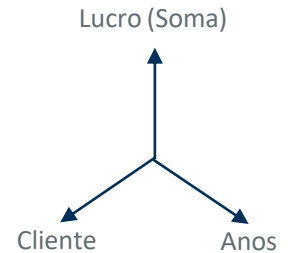


# IDENTIFICANDO INDICADORES E DIMENSÕES

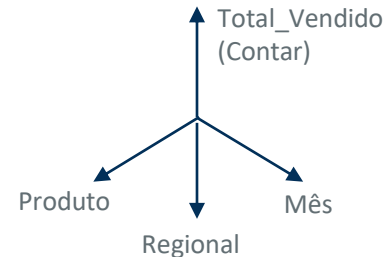


# IDENTIFICANDO INDICADORES E DIMENSÕES

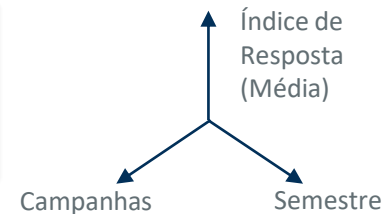
**Fato 1:** Qual foi o cliente mais lucrativo da empresa ao longo dos anos?



**Fato 2:** Qual o produto mais vendido por regional em cada mês?



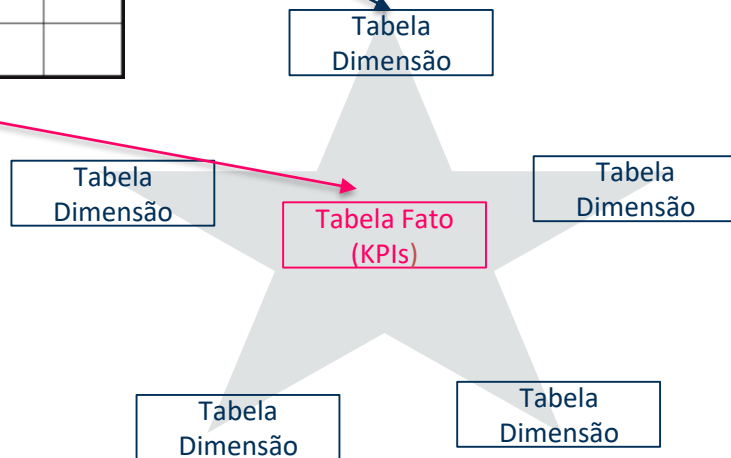
**Fato 3:** Qual o índice de resposta médio trazido pelas campanhas feitas no segundo semestre?



# IDENTIFICANDO INDICADORES E DIMENSÕES

## Business Process Matrix (BPM) – Star Schema

Processo	Indicador	Dim. Data	Dim 2	Dim 3	Dim 4	Dim 5	Dim 6
Proc 1	KPI 1	x	x	x		x	x
Proc 2	KPI 2	x	x		x	x	
Proc 3	KPI 3	x					x
Proc 4	KPI 4	x		x			





## EXEMPLO – USO DO BPM

**Fato 1:** Qual é o total de Unidades Disponíveis e Unidades Vendidas dentro das torres de uma série de empreendimentos?

**Fato 2:** Fato 2 -- > Qual são as Unidades mais visualizadas no site?

<b>Modelo:</b>	Star Schema			
<b>Grão:</b>	Unidade / Data da Venda ou Lançamento			
<b>Métrica</b>	<b>Tipo de Agregação</b>	<b>Tipo de Origem</b>		
Total de Unidades Disponíveis	Soma/Média	Quente / Estruturada		
Total de Unidades Vendidas	Soma/Média	Quente / Estruturada		
Total de Unidades Visualizadas	Soma/Média	Quente / Semiestruturada		
<b>Dimensão</b>	<b>Tipo de Dimensão</b>			
Unidade	Hierárquica			
Torre				
Empreendimento				
Data	Tempo			
<b>Matriz Métrica/Dimensão</b>	Unidade	Torre	Empreendimento	Data
Total de Unidades Disponíveis	x	x	x	x
Total de Unidades Vendidas	x	x	x	x
Total de Unidades Visualizadas	x	x	x	x

# VAMOS APLICAR?

1. Escolha sua dupla.
2. Leia e discuta com o seu par o paper – “Projeto - Trade Marketing”.
3. Faça o levantamento das necessidades apresentadas no Case.
4. Considere a necessidade de criar quatro indicadores:
  - Preço do produto no PDV
  - Quantidade Distribuída do produto para o PDV
  - Quantidade Vendida do produto no PDV
  - Quantidade de Visualizações por produto no site

Especifique os requisitos multidimensionais construindo a matriz métrica/dimensão (BPM). Faça duas matrizes, uma por canal.

*\*Importante: Insira o resultado do Hands On (3 e 4) no item 2 de “Requisitos Funcionais” do documento – “Especificação da Solução de Dados”.*

*Para conhecer os dados do canal do varejo físico abra o arquivo “Amostra de Dados - Trade”, e para o e-commerce abra o arquivo de Log.*

# VAMOS APLICAR?

Teste seus conhecimentos!

*[Clique aqui e responda.](#)*

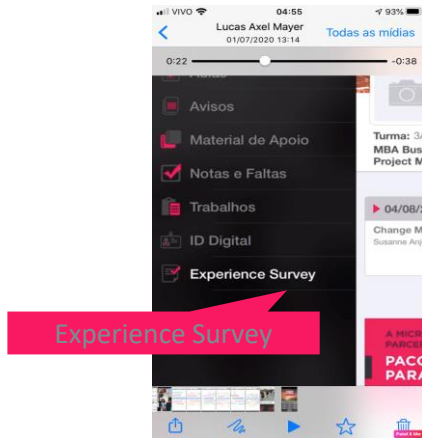


# Como foi a sua experiência com a da aula de hoje?

FIAP MBA+

## Pelo aplicativo da FIAP

(Entrar no FIAPP, e no menu clicar em Experience Survey)



## Pela Web



# Hoje foi o início da nossa jornada;)



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



[profassiana.campos@fiap.com.br](mailto:profassiana.campos@fiap.com.br)

FIAP

Copyright © 2022 | Professora Tassiana Rugoni de Campos  
Todos os direitos reservados. Reprodução ou divulgação total ou parcial deste documento, é expressamente proibido sem consentimento formal, por escrito, do professor/autor.

