

## **Data Science em uma Estação de Tratamento de Esgoto**

### *Data Science in a Sewage Treatment Plant*

**ALLAN DOS ANJOS PESTANA**

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE SÃO PAULO - IFSP

**LINCOLN LUAN SOUZA TELES**

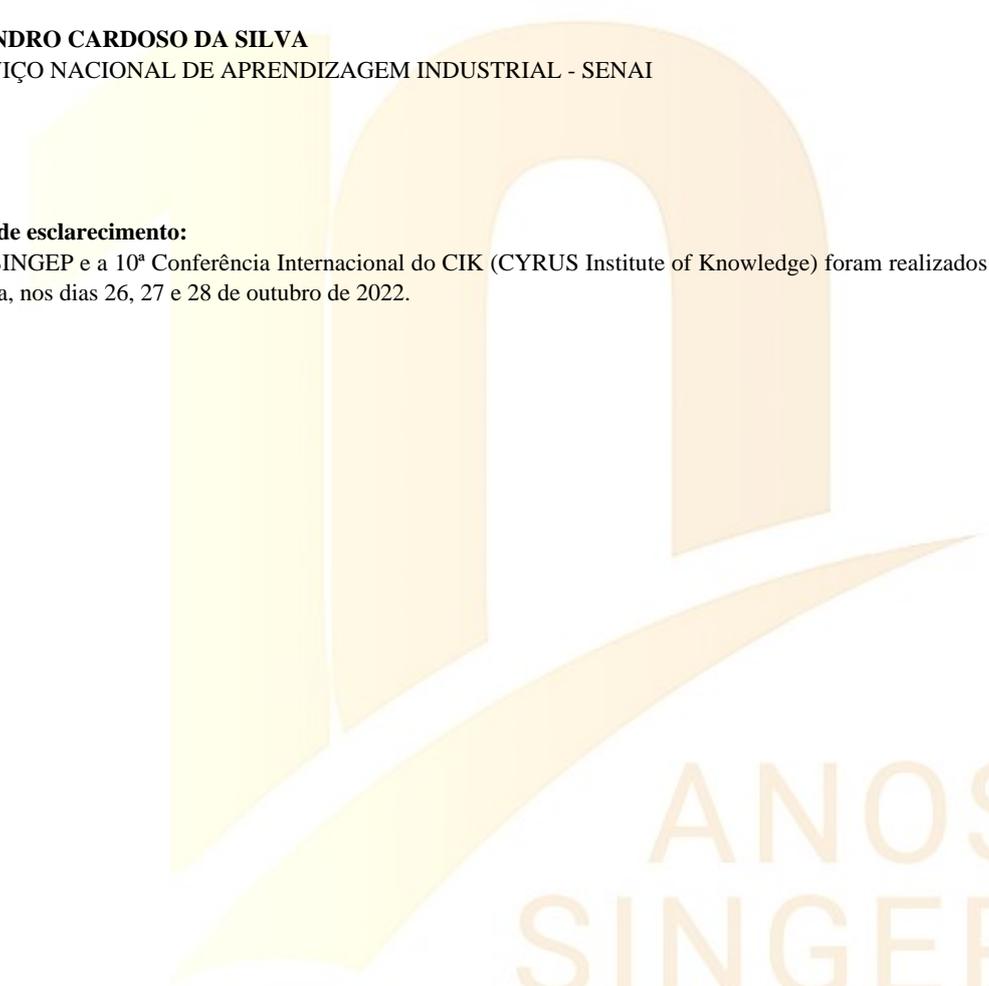
SERVIÇO NACIONAL DE APRENDIZAGEM INDUSTRIAL - SENAI

**LEANDRO CARDOSO DA SILVA**

SERVIÇO NACIONAL DE APRENDIZAGEM INDUSTRIAL - SENAI

#### **Nota de esclarecimento:**

O X SINGEP e a 10ª Conferência Internacional do CIK (CYRUS Institute of Knowledge) foram realizados de forma remota, nos dias 26, 27 e 28 de outubro de 2022.



ANOS  
SINGEP

## **Data Science em uma Estação de Tratamento de Esgoto**

### **Objetivo do estudo**

O objetivo é apresentar um piloto e os possíveis benefícios do uso das tecnologias habilitadoras da Indústria 4.0 no saneamento, mais especificamente em uma Estação de Tratamento de Esgotos (ETE).

### **Relevância/originalidade**

Uso de Data Science em uma Estação de Tratamento de Esgoto com um nível elevado de instrumentação analítica visando otimização e predição e consequentemente eficiência energética e operacional da planta.

### **Metodologia/abordagem**

Piloto de Data Science nas Estações estudadas com o uso de um add-in para excell e observação das correlações e resultados obtidos.

### **Principais resultados**

Altos níveis de correlação observados, e potencial de redução de consumo energético e melhoria na eficiência operacional da Estação.

### **Contribuições teóricas/metodológicas**

Avanço do Data Science no Saneamento.

### **Contribuições sociais/para a gestão**

Levando-se em consideração que as Estações de Tratamento de Esgoto, retornam um bem finito ao meio ambiente, maior performance das mesmas trazem benefícios diretos e indiretos a sociedade e ao meio ambiente.

**Palavras-chave:** Automação, Indústria 4.0, Tratamento de Esgoto, Data Science

ANOS  
SINGEP

## *Data Science in a Sewage Treatment Plant*

### **Study purpose**

The objective is to present a pilot and the possible benefits of using the enabling technologies of Industry 4.0 in sanitation, more specifically in a Sewage Treatment Station (ETE).

### **Relevance / originality**

Use of Data Science in a Sewage Treatment Plant with a high level of analytical instrumentation aiming at optimization and prediction and consequently energy and operational efficiency of the plant.

### **Methodology / approach**

Data Science pilot in the studied stations with the use of an add-in for excell and observation of the correlations and results obtained.

### **Main results**

High levels of correlation observed, and potential for energy consumption reduction and improvement in the operational efficiency of the Station.

### **Theoretical / methodological contributions**

Advancement of Data Science in Sanitation.

### **Social / management contributions**

Taking into account that the Sewage Treatment Stations return a finite good to the environment, their greater performance brings direct and indirect benefits to society and the environment.

**Keywords:** Automation, Industry 4.0, Sewage Treatment, Data Science

ANOS  
SINGEP

## 1 Introdução

O objetivo é apresentar um piloto e os possíveis benefícios do uso das tecnologias habilitadoras da Indústria 4.0 no saneamento, mais especificamente em uma Estação de Tratamento de Esgotos (ETE), visando, por exemplo, melhoria no controle de Oxigênio Dissolvido (OD) nos tanques de aeração, controle dos produtos químicos utilizados na etapa de desidratação mecânica de lodo, controle do manto de lodo nos decantadores primários e secundários, entre muitos outros e consequentemente a estabilidade geral nos processos subsequentes. Visa também a eficiência energética nestas etapas do processo, considerando a Companhia de Saneamento, um dos maiores clientes da Companhia de Energia Elétrica da Região, e alguns de seus equipamentos podendo ser comparados ao consumo energético de pequenos municípios.

A Estação de Tratamento de Esgotos atende diversas cidades dessa região, e tem vazão média de 2792 litros/segundo e beneficia mais de 1,4 milhões de habitantes.

Essa Estação tem concepção de tratamento por lodo ativado, um tratamento aeróbico, onde é enviado oxigênio ao licor misto nos tanques de aeração. Dessa forma, os microrganismos presentes neste licor decompõem a matéria orgânica presentes no esgoto, segue fluxograma de uma dessas Estações na figura 1:

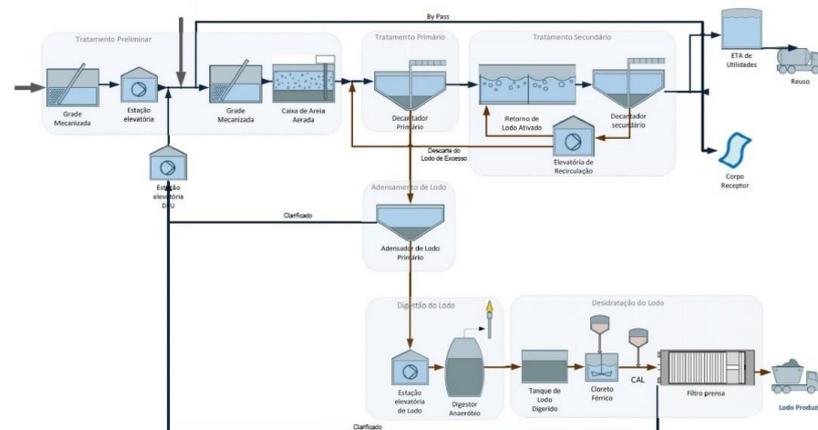


Figura 1 – Fluxograma do tratamento – ETE  
Fonte: O autor (2022).

Segundo Van Haandel e Marais (1999) para garantir que o tratamento de esgoto seja eficiente é preciso manter um valor mínimo de Oxigênio Dissolvido (OD) no licor misto (de 1 a 2 mg/L).

A automação e controles simples podem auxiliar nesse controle, que é um exemplo, dentre os diversos controles existentes na planta.

Outro fato observado, é que a carga orgânica da estação pode ser muito variável, dependendo de diversos fatores, como por exemplo, horário do dia, período de chuvas, entre outros. A variação desses fatores, exige que a oxigenação nos tanques de aeração varie imediatamente. Segundo Metcalf e Eddy (2003) os aeradores devem permitir atender essas variações com flexibilidade e economizar energia, dado o custo energia nessa etapa do

processo. Jordão e Pessoa (2014) afirmam que o oxigênio dissolvido é um parâmetro fundamental no tratamento por lodo ativado para decomposição da matéria orgânica.

## 2 Referencial Teórico

### 2.1 Pirâmide de Automação

A literatura de automação, divide os elementos pertencentes a um sistema automático em cinco camadas, visando separá-los de acordo com sua funcionalidade hierárquica. A primeira camada, contempla todos os dispositivos de campo fundamentais para os controles, sejam atuadores, medidores de nível, vazão, pH e afins. Na segunda camada da pirâmide estão os controladores de processo, no caso das plantas de tratamento, são utilizados majoritariamente os Controladores Lógicos Programáveis (CLPs). Na terceira camada estão os sistemas de supervisão e as Interface Homem Máquina (IHMs) com a funcionalidade de realizar a primeira interface com usuário final. Na quarta camada, os gerenciadores da planta, e na Companhia de Saneamento em questão é utilizado o Plant Information Management System - Sistema de gerenciamento de informações da planta (PIMS), PI System da OSI Software. Por fim, na última camada estão os Enterprise Resource Planning - Sistema integrado de gestão empresarial (ERPs) e atualmente o utilizado na Companhia é o SAP (Goeking, 2010).

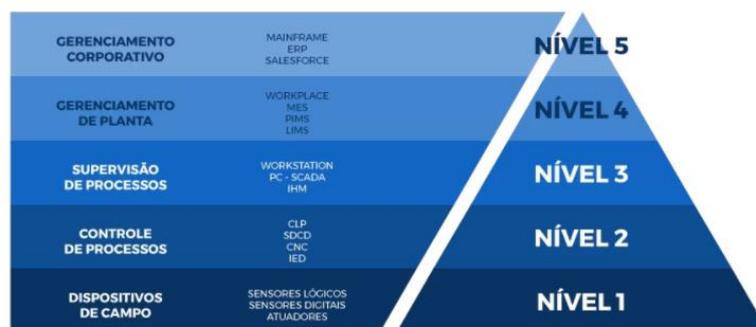


Figura 2 – Pirâmide de automação

Fonte: <https://www.altus.com.br/post/100/curso-de-introducao-a-automacao--5baula-01-5d>

### 2.2 IoT

Internet of things ou Internet das coisas, é um conceito fundamental para a indústria, onde o “caminho” de comunicação usualmente utilizado por seres humanos na internet, para realizar compras, se comunicar em redes sociais e afins, é nesse caso utilizado pelas máquinas (Sacomano et al., 2018).

### 2.3 Big Data

O termo ou conceito Big Data não é exatamente novo. De 2004 a 2010, por exemplo, a busca pelo tema se mantinha constante, foi então que o interesse pelo assunto começou a aumentar exponencialmente, chegando a ser dez vezes maior em 2018 (Google Trends, 2022). E como é possível afirmar isso sem muitas pesquisas? Utilizando a ferramenta do Google chamada Google Trends, que a partir de um número gigantesco de buscas em sua plataforma, converte esses dados em gráficos que relacionam as pesquisas em determinados locais e períodos do tempo, um exemplo fantástico de Big Data.

Um outro exemplo de Big Data pode ser visto nos mapas virtuais. Antes o mapa do GPS, era egoísta, encaminhava todos pelo “melhor caminho” que a partir daí se abarrotava de motoristas e já não se mostrava ser exatamente o melhor. Hoje o Waze é colaborativo, utilizando-se dos dados dos usuários, indica o melhor caminho para cada indivíduo, e quando aquele caminho se tornar saturado de motoristas, o segundo melhor passará a ser a melhor rota, tudo isso de forma dinâmica. E isso só é possível com a utilização de Big Data.

Um startup Recifeense, Epitrack, entre os finalistas do The Venture Brasil, etapa brasileira de uma importante competição de empreendedorismo social no mundo, promete utilizar de Big Data, para prever com até duas semanas de antecedência o aparecimento de alguma epidemia ou surto de doença, através das pesquisas realizadas no Google pelas pessoas. Em sua definição, Big Data, é uma grande quantidade de dados em volume, variedade e velocidade segundo Gartner (2022). Nesse caso, uma grande quantidade de pessoas pesquisando, por exemplo, “manchas vermelhas”, poderia significar o surto de determinada doença, em algum local, antes mesmo dos seres humanos especialistas no assunto perceberem isso. A máquina fazendo o que o ser humano não poderia fazer sozinho.

A era da informação permitiu a existência de uma grande quantidade de dados, com infinitas possibilidades de utilização, mas segundo Stair e Reynolds (2013) dados somente não é informação, e a boa utilização ou não de todo esse material, pode significar o surgimento de uma nova era ainda desconhecida para todos nós, ou fracasso de grandes companhias sem visão de negócio.

## **2.4 Indústria 4.0**

Toda evolução humana e tecnológica que aconteceu gradualmente nos últimos séculos trouxeram benefícios para a indústria e meios de produção. A chamada primeira revolução industrial ocorre na segunda metade do século XVIII com a criação da máquina a vapor e a diminuição significativa do esforço braçal humano no trabalho. A segunda através da produção em escala e divisão do trabalho com o advento da eletricidade no começo do século XX. A terceira revolução é tida no início dos anos 70, com os robôs e controladores programáveis. (Kagermann; Wahlster; Helbig, 2013).

A quarta revolução industrial é caracterizada pelos sistemas autônomos e capazes de tomar ou propor decisões até então impossíveis ao ser humano, haja vista a capacidade de processamento. Segundo Choi, Kang, Jung, Kulvatunyou e Morris (2016) é também caracterizado pela utilização de sistemas inteligentes e elevado grau de automação. O termo foi utilizado pela primeira vez em 2011 em Hanover na Alemanha segundo Drath e Horch, (2014).

## **2.5 Saneamento 4.0**

Assim como a indústria, o saneamento demanda processos cada vez melhores e mais eficientes. Ainda levando em consideração o novo marco legal do saneamento, lei nº 14.026, de 15 de julho de 2020, que preconiza o setor de saneamento atender 99% da população com água potável e 90% com coleta e tratamento de esgoto até 31 de dezembro de 2033, o setor de saneamento tem um grande desafio pela frente.

O Saneamento 4.0 se mostra como uma alternativa para o setor. Alguns dos pilares da indústria 4.0, tem a possibilidade de tornar, por exemplo, o tratamento de esgoto algo não mais

reativo as mudanças de características do esgoto, do clima ou das bacias a que atende e sim podendo prever essas alterações com o uso de Big Data e Machine Learning.

Muito controles necessários nas Estações de Tratamento de Esgoto, como por exemplo, controle de oxigênio dissolvido nos tanques de aeração, cloreto férrico e polímero na desidratação mecânica de lodo, ou hipoclorito na água de reuso já existem desde o início dessas Estações. Contudo, com o Saneamento 4.0, espera-se aplicar as tecnologias habilitadoras da Indústria 4.0 nesse setor, para alcançar um patamar que o controle humano não permite. Por exemplo, utilizar do Big Data, para minerar grande quantidade de dados relacionados as bacias pertencentes a determinada Estação, fatores climáticos e externos para prever e antecipar a atuação dos controles, tornando os processos muito mais eficientes.

## 2.6 Machine Learning

O Machine Learning pode ser definido como a capacidade de uma máquina de aprender sem ser previamente programado (Samuel, 1959). Posteriormente redefinido e categorizado em quatro possíveis grupos de algoritmos de Machine Learning: supervisionado, não supervisionado, semisupervisionado e aprendizado por reforço (Mitchell, 1997).

## 2.7 Inteligência Artificial

A inteligência artificial tem por objetivo fazer com que determinada máquina se comporte ou pense com inteligência segundo Gomes (2010). Por exemplo, em um dia de chuva, é comum que uma pessoa leve um guarda-chuva, em um dia de chuva e frio é esperado também que essa mesma pessoa leve um guarda-chuva e uma blusa. Agora imagine em uma indústria em que diversos fatores acarretam uma porção de implicações. É bem provável que a capacidade humana, não seja suficiente para dar a esses processos todas a resposta de que ele precise, e as máquinas, sua capacidade computacional e a inteligência artificial podem ser uma alternativa para tal.

## 2.8 Avaliação de Maturidade da Acatech

Para identificação de qual momento a empresa está na implementação das tecnologias habilitadoras da indústria 4.0, é possível utilizar uma avaliação de maturidade, como por exemplo, a da ACATECH (German Academy of Science and Engineering – Academia Alemã de Ciências e Engenharia) de Schumacher, Erol e Sihm (2016) e Schuh et al. (2017).

Como apontado em algumas revisões bibliográficas, a Europa é o local, segundo suas publicações, com maior avanço no uso das tecnologias habilitadoras atualmente. No Brasil, infelizmente, boa parte da indústria ainda se encontra entre a segunda e terceira revolução industrial, sendo o setor automotivo o que apresenta melhor desempenho nesse sentido (FIRJAN, 2016).



Figura 3 – Escala de maturidade – Indústria 4.0 – ACATECH  
Fonte: <https://www.altus.com.br/post/100/curso-de-introducao-a-automacao--5baula-01-5d>

### 3 Metodologia

Trabalhos pontuais de executados anteriormente no departamento demonstram os resultados financeiros e operacionais obtidos com o investimento em automação nos processos. Contudo o potencial de benefícios projetados pode ir muito além. A Unidade de Tratamento em questão que, por exemplo, possuía poucos instrumentos analíticos para controle de seus processos, tem previsão de aumento exponencial dessas análises, conforme pode ser visto abaixo:

Tabela 1: Instrumentação analítica em processo de comissionamento 03/2022

	Sistema	Parâmetro	Quant.
Compra de dados	Sistema 1	pH	5
		Condutividade	5
		SST	6
		DQO	2
		ORP	6
		OD	12
		NH3	3
		Turbidez	6
		Manto de lodo	6
		<b>Total</b>	<b>51</b>
	Sistema 2	pH	1
		Condutividade	2
		SST	7
		DQO	2
		ORP	12
		OD	12
		NH3	3
		Turbidez	2
		Manto de lodo	4
		Cloro	1
	<b>Total</b>	<b>46</b>	
	Sistema 3	pH	1
		Condutividade	2
		SST	4
		DQO	2
ORP		3	
OD		3	
NH3		3	
Turbidez		2	
Manto de lodo		4	
Cloro		1	
<b>Total</b>	<b>25</b>		

Fonte: Do autor, 2022.

A previsão de liberação desses dados é para fevereiro de 2022, e após poderá ser pilotada uma solução de Machine Learning para o Tratamento de Esgoto, nesse caso, pela quantidade de análises, pioneira no Brasil. Em conjunto com o contrato de compra de dados também está em andamento um contrato para compra de elementos finais de controle, afim de que os mesmos possam realizar atuação no processo, podendo ser possível fechar as malhas de controles em cada uma das Estações de Tratamento.

Tabela 2 – Elementos finais de controle

Contrato	Sistema/ETE	Objeto	Etapa do processo	Quantidade
Elementos finais de controle	ETE 1	Transmissor de nível radar	Poço de espuma	4
		Transmissor termal mássico de ar	Tanque de aeração	4
		Válvula + atuador	Decantador primário	18
		Caixa de terminais de válvulas	Decantador primário	2



ETE 2	Compressor de ar	Decantador primário	2
	Transmissor termal mássico de ar	Tanque de aeração	12
	Compressor de ar	Decantador primário	2
	Atuador + posicionador	Decantador primário	4
ETE 3	Atuador + posicionador	Tanque de aeração	12
	Válvula + atuador	Decantador primário	12
	Caixa de terminais de válvulas	Decantador primário	1
	Compressor de ar	Decantador primário	2
	Atuador + posicionador	Tanque de aeração	4

Fonte: Do autor, 2022.

Com os dados analíticos das plantas de tratamento de esgoto tabulados, estudamos a possibilidade de haver correlações entre as séries de dados. O Excel possui uma fórmula para o cálculo do coeficiente de correlação, CORREL. Abaixo a equação que a fórmula utiliza:

$$Correl(X, Y) = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

Equação (1)

Ou seja, a partir desse ponto podemos enxergar possíveis correlações entre dados analíticos de diferentes etapas do processo. Como esses dados ainda estão em processo de registro segue exemplo da planilha criada no Excel, afim de criar um piloto para o estudo, inicialmente com dados aleatórios, com periodicidade de 5 minutos:

Data	Hora	TAG	Local		Parâmetro
			Esgoto bruto	Esgoto tratado	
01/04/2022	00:00:00	P1-MT-101	pH	Esgoto bruto	0,98
01/04/2022	03:00:00	P1-MT-102	Condutividade	Esgoto bruto	0,88
01/04/2022	06:00:00	P1-MT-103	SST (merask)	Esgoto tratado	0,85
01/04/2022	09:00:00	P1-MT-104	DOO	Esgoto tratado	0,42
01/04/2022	12:00:00	P1-MT-105	NH3	Captação Aquático	0,77
01/04/2022	15:00:00	P1-MT-106	SST (merask)	Etanque Primário	0,48
		P1-MT-107	DOO	Etanque Final	0,53
		P1-MT-108	SST (merask)	Litor misto	0,54
		P1-MT-109	SST (merask)	Litor misto	0,55
		P1-MT-110	ORP	Reator Aéreo 1A	0,66
		P1-MT-111	ORP	Reator Aéreo 1B	0,68
		P1-MT-112	ORP	Reator Aéreo 2A	0,66
		P1-MT-113	ORP	Reator Aéreo 2B	0,68
		P1-MT-114	ORP	Reator Aéreo 3A	0,66
		P1-MT-115	ORP	Reator Aéreo 3B	0,68
		P1-MT-116	Oxigênio Dissolvido	Reator Aéreo 1A	0,50
		P1-MT-117	Oxigênio Dissolvido	Reator Aéreo 1B	0,51
		P1-MT-118	Oxigênio Dissolvido	Reator Aéreo 2A	0,50
		P1-MT-119	Oxigênio Dissolvido	Reator Aéreo 2B	0,51
		P1-MT-120	Oxigênio Dissolvido	Reator Aéreo 3A	0,50
		P1-MT-121	Oxigênio Dissolvido	Reator Aéreo 3B	0,51
		P1-MT-122	Oxigênio Dissolvido	Reator Aéreo 2A	0,50
		P1-MT-123	Oxigênio Dissolvido	Reator Aéreo 2B	0,51
		P1-MT-124	Oxigênio Dissolvido	Reator Aéreo 3A	0,50
		P1-MT-125	Oxigênio Dissolvido	Reator Aéreo 3B	0,51

Figura 4 – Dados analíticos (ETE 1)

Fonte: Do autor, 2022.

Como em alguns pontos, existe mais de uma medição de determinada grandeza, como por exemplo, turbidez nos DSs, e não há interesse em enxergar correlações entre essas medidas individualmente, então nesses casos a média dos valores encontrados devem servir para os passos seguinte. Abaixo segue a tabela, com o detalhe em cinza para os dados em que foram obtidas as médias:



Data	Hora	TAG	Parâmetro	Local
01/04/2022	00:00:00	P1-AIT-101	pH	Esgoto bruto
01/04/2022	03:00:00	P1-AIT-102	Condutividade	Esgoto bruto
01/04/2022	06:00:00	P1-AIT-103	SST (imersão)	Esgoto bruto
01/04/2022	09:00:00	P1-AIT-104	DOO	Esgoto desarenado
01/04/2022	12:00:00	P1-AIT-105	NH3	Esgoto desarenado
		P1-AIT-123	SST (imersão)	Captação Aquapolo
		P1-AIT-105	SST (imersão)	Efluente Primário
		P1-AIT-506	DOO	Efluente Final
		P1-AIT-406	SST (imersão)	Licor misto
		P1-AIT-408	ORP (Médio) - Saída	Reatores Aeróbios
		P1-AIT-414	Oxigênio Dissolvido (Médio) - Entrada	Reatores Aeróbios
		P1-AIT-420	Oxigênio Dissolvido (Médio) - Saída	Reatores Aeróbios
		P1-AIT-426	NH3	Efluente Primário
		P1-AIT-427	NH3	Efluente Final
		P1-AIT-428	Turbidez (Média)	Efluente DSs
		P1-LIT-534	Mano de lodo (Médio)	Tanque DSs
		P1-AIT-440	pH	Efluente Final
		P1-AIT-601	SST (imersão)	Lodo Alu. Digestor
		P1-AIT-602	SST (imersão)	Lodo condicionado

Figura 5 – Dados analíticos, média (ETE 1)  
Fonte: Do autor, 2022.

E então é possível calcular o coeficiente de correlação entre duas séries de dados. Abaixo temos uma tabela com todas as correlações sendo calculadas, o valor pode ser obtido cruzando linha e coluna com os dados em que se deseja correlacionar:

				Comparação Real time																			
				Local																			
				Parâmetro																			
				TAG																			
Local	Parâmetro	TAG		P1-AIT-101	P1-AIT-102	P1-AIT-103	P1-AIT-104	P1-AIT-123	P1-AIT-105	P1-AIT-506	P1-AIT-406	P1-AIT-408	P1-AIT-414	P1-AIT-420	P1-AIT-426	P1-AIT-427	P1-AIT-428	P1-LIT-534					
Esgoto bruto	pH	P1-AIT-101			0,48	-0,06	0,24	-0,39	0,34	0,00	-0,40	0,25	-0,72	0,38	-0,15	-0,31	0,70	0,00	-0,00	-0,00			
Esgoto bruto	Condutividade	P1-AIT-102				0,22	0,22	-0,29	0,15	0,55	-0,62	0,04	-0,79	0,20	0,26	-0,05	0,69	0,01	-0,01	-0,01			
Esgoto desarenado	SST (imersão)	P1-AIT-103					0,71	-0,44	0,31	0,41	0,04	0,38	0,15	-0,17	-0,26	-0,27	-0,34	0,33	-0,00	-0,00			
Esgoto desarenado	DOO	P1-AIT-104						-0,24	0,18	0,12	0,12	-0,15	0,14	0,48	0,09	-0,21	-0,03	0,70	0,00	-0,00			
Captação Aquapolo	NH3	P1-AIT-123							-0,95	0,02	0,37	-0,65	0,20	-0,02	0,35	0,32	0,15	0,07	0,00	0,00			
Efluente Primário	SST (imersão)	P1-AIT-105								-0,14	-0,22	0,57	-0,15	0,15	-0,27	-0,38	-0,19	-0,17	0,00	0,00			
Efluente Final	DOO	P1-AIT-506									-0,09	0,24	-0,31	-0,40	0,26	-0,08	0,35	0,21	-0,01	-0,01			
Licor misto	SST (imersão)	P1-AIT-406										-0,08	0,63	0,09	0,35	0,34	-0,45	0,22	0,00	0,00			
Reatores Aeróbios	ORP (Médio) - Saída	P1-AIT-408												-0,19	-0,57	-0,53	-0,18	-0,19	-0,34	-0,00			
Reatores Aeróbios	Oxigênio Dissolvido (Médio) - Entrada	P1-AIT-414													-0,12	0,02	0,26	-0,64	0,42	0,00			
Reatores Aeróbios	Oxigênio Dissolvido (Médio) - Saída	P1-AIT-420														0,50	0,05	0,30	0,28	0,00			
Efluente Primário	NH3	P1-AIT-426														0,55	0,30	0,36	0,00	0,00			
Efluente Final	NH3	P1-AIT-427																-0,12	0,19	0,00			
Efluente DSs	Turbidez (Média)	P1-AIT-428																	-0,01	-0,00			

Figura 6 – Dados analíticos, correlações (ETE 1)  
Fonte: Do autor, 2022.

Fonte		Alinhamento																			
=CORREL(\$B\$5:\$B\$13;\$D\$5:\$D\$13)																					
AT	AV	AW	AX	AY	AZ	BB	BC	BD	BE	BF	BG	BH	BI	BJ	BK	BL	BM	BN	BO	BP	
Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	
SST (imersão)	SST (imersão)	SST (imersão)	NH3	Condutividade																	
PH-RT-001	PH-RT-002	PH-RT-003	PH-RT-004	PH-RT-005	Data	Hora	TAG	PH-RT-001	PH-RT-002	PH-RT-003	PH-RT-004	PH-RT-005	PH-RT-006	PH-RT-007	PH-RT-008	PH-RT-009	PH-RT-010	PH-RT-011	PH-RT-012	PH-RT-013	
0,21	0,76	0,62	0,17	0,50	01/04/2022	00:00:00	0	0,26	0,38	0,10	0,37	0,09	0,85	0,32	0,35	0,17	0,52	0,53	0,73	0,72	0,4
0,71	0,83	0,34	0,27	0,43	01/04/2022	03:00:00	0	0,04	0,66	0,22	0,63	0,38	0,08	0,13	0,46	0,63	0,43	0,46	0,32	0,22	0,7
0,75	0,38	0,66	0,57	0,30	01/04/2022	06:00:00	0	0,50	0,36	0,19	0,50	0,50	0,72	0,75	0,50	0,29	0,39	0,32	0,10	0,93	0,6
0,52	0,71	0,34	0,83	0,20	01/04/2022	09:00:00	0	0,26	0,75	0,42	0,64	0,58	0,31	0,85	0,10	0,56	0,42	0,40	0,31	0,02	0,5
0,85	0,30	0,83	0,71	0,58	01/04/2022	12:00:00	0	0,34	0,83	0,03	0,88	0,60	0,75	0,43	0,85	0,51	0,35	0,69	0,64	0,78	0,3
0,68	0,76	0,51	0,67	0,23	01/04/2022	15:00:00	0	0,79	0,30	0,68	0,10	0,05	0,35	0,46	0,20	0,42	0,47	0,66	0,82	0,47	0,6
0,64	0,58	0,38	0,45	0,11	01/04/2022	18:00:00	0	0,55	0,60	0,67	0,81	0,53	0,25	0,02	0,61	0,46	0,43	0,36	0,08	0,95	0,3
0,75	0,67	0,68	0,04	0,44	01/04/2022	21:00:00	0	0,59	0,56	0,21	0,73	0,71	0,09	0,16	0,53	0,48	0,43	0,47	0,22	0,21	0,4
0,45	0,79	0,35	0,15	0,78	01/04/2022	00:00:00	0	0,32	0,34	0,15	0,54	0,93	0,22	0,14	0,26	0,64	0,60	0,49	0,46	0,20	0,4

Figura 7 – Correlações, seleção da série de dados  
Fonte: Do autor, 2022.

Contudo, como de conhecimento de estudiosos da área de tratamento de esgoto, alguns dados podem se correlacionar com outros, porém com determinado “delay”. Por exemplo, é de conhecimento que o Oxigênio Dissolvido nos Tanques de aeração influencia na Amônia na etapa secundária, porém com o atraso de algumas horas devido ao tempo de detenção hidráulica dessa etapa do processo, ou seja, o tempo que aquela amostra analisada demora até chegar ao próximo analisador. Então, abaixo segue análise dos mesmos dados, porém considerando esse delay:

Local	Parâmetro	TAG	Comparação Delay 3 horas																	
			PH-RT-001	PH-RT-002	PH-RT-003	PH-RT-004	PH-RT-005	PH-RT-006	PH-RT-007	PH-RT-008	PH-RT-009	PH-RT-010	PH-RT-011	PH-RT-012	PH-RT-013	PH-RT-014	PH-RT-015	PH-RT-016	PH-RT-017	
Esgoto bruto	pH	PH-AIT-01		-0,84	0,49	0,29	0,29	-0,59	-0,53	-0,07	0,34	0,24	-0,16	-0,34	-0,17	-0,06	-0,08	0,36		
Esgoto bruto	Condutividade	PH-AIT-02			0,57	-0,04	-0,85	-0,18	0,13	-0,24	0,15	-0,20	0,01	0,38	-0,07	0,27	-0,38	0,71		
Esgoto desarenado	SST (imersão)	PH-AIT-03				0,67	0,44	-0,11	-0,45	0,65	-0,14	-0,33	-0,21	-0,63	0,32	-0,77	-0,45	-0,28		
Esgoto desarenado	DOO	PH-AIT-04					-0,22	-0,13	0,06	-0,19	0,17	0,27	0,42	0,67	-0,46	0,55	0,07	-0,31		
Captação Aquapolo	NH3	PH-AIT-123						0,25	0,33	-0,27	0,06	0,16	0,53	0,11	-0,33	-0,13	0,50	-0,50		
Efluente Primário	SST (imersão)	PH-AIT-105							0,11	-0,41	0,32	-0,01	0,23	0,64	-0,43	0,42	0,31	0,58		
Efluente Final	DOO	PH-AIT-506								0,07	0,18	-0,46	0,41	0,24	0,09	-0,32	-0,12	0,30		
Lícor misto	SST (imersão)	PH-AIT-406									-0,10	0,45	0,11	0,16	-0,49	0,48	0,28	-0,14		
Reatores Aeróbios	ORP (Médio) - Saída	PH-AIT-408										-0,28	0,19	-0,36	0,62	-0,31	0,08	-0,80		
Reatores Aeróbios	Oxigênio Dissolvido (Médio) - Entrada	PH-AIT-414											-0,57	-0,14	0,25	0,24	-0,12	0,18		
Reatores Aeróbios	Oxigênio Dissolvido (Médio) - Saída	PH-AIT-420												0,27	0,43	0,08	-0,24	0,13		
Efluente Primário	NH3	PH-AIT-426													0,69	0,35	-0,28	0,11		
Efluente Final	NH3	PH-AIT-427														0,34	-0,30	0,33		
Efluente DSs	Turbidez (Média)	PH-AIT-428															0,13	0,30		

Figura 8 – Dados analíticos, correlações com delay 1 hora (ETE 1)  
Fonte: Do autor, 2022.



<b>22</b>	0,02	-0,05	0,07	0,01	-0,01	0,03	-0,04	0,07	-0,06	0,01	-0,05	0,12	-0,03	0,10	-0,05	0,04	0,06	-0,06	-0,01	-0,1
<b>23</b>	0,06	-0,02	0,05	-0,03	-0,09	-0,03	-0,11	0,07	0,02	-0,04	0,05	0,02	0,07	-0,14	-0,04	0,04	-0,02	0,05	0,04	-0,1
<b>24</b>	-0,03	0,04	-0,01	0,02	-0,01	-0,04	0,02	0,02	-0,01	-0,02	0,02	-0,04	-0,05	-0,09	-0,03	0,03	-0,09	0,04	0,02	-0,1
<b>SOMA</b>	0,19	-0,44	0,19	0,00	-0,21	-0,07	-0,43	0,06	0,06	-0,04	0,34	-0,11	-0,01	-0,20	-0,13	0,13	0,12	-0,09	0,06	0,0

Figura 11 – Soma das correlações  
 Fonte: Do autor, 2022.

Por fim, é possível visualizar os gráficos para cada uma das correlações, com base nessa “biblioteca” posteriormente criar padrões de operação e para os próprios controladores responsáveis por esses processos:

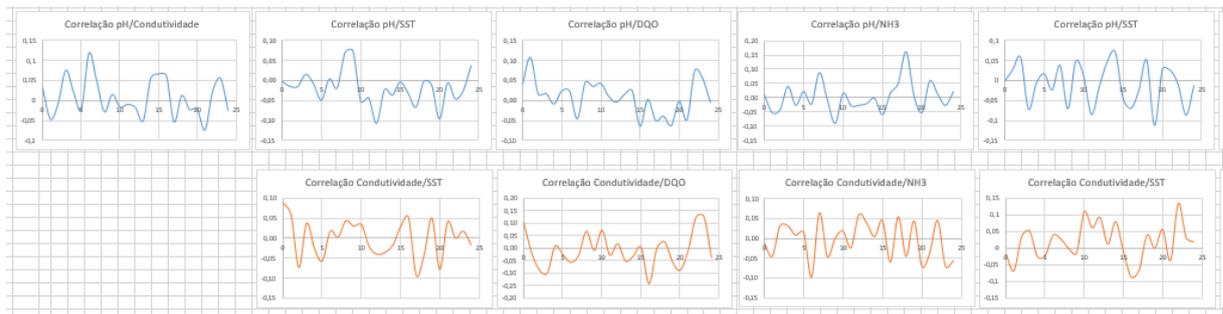


Figura 12 – Gráficos da modelagem  
 Fonte: Do autor, 2022.

#### 4 Análise e discussão dos resultados

Nessa etapa é possível já simular o uso da planilha, inicialmente de maneira manual, com dados reais, por exemplo, da ETE 1, e foi possível enxergar algumas correlações inclusive considerando o tempo de detenção hidráulica das etapas envolvidas:

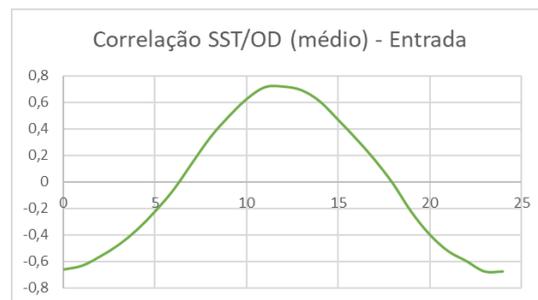


Figura 13 – Correlação do SST com OD em diversos “delays”  
 Fonte: Do autor, 2022.

Com esse gráfico podemos concluir que há uma forte correlação (0,72) entre os dados de Sólidos Suspensos Totais no esgoto desarenado com o Oxigênio Dissolvido nos Tanques de Aeração na ETE 1, e não somente isso, mas conseguimos concluir que essa correlação é maior depois de exatamente 11 horas. Contudo, nesse caso, é mais coerente observar a correlação vista na mesma base de tempo (-0,66), inversamente proporcional, haja vista que o tempo de detenção hidráulica entre essas etapas é pequeno, e que a correlação entre essas análises tende a ser contrária, ou seja, quanto maior o SST, maior a carga que está sendo enviada aos tanques, diminuindo assim, o valor observado de oxigênio dissolvido.

Essa correlação, agora observando-se os dados propriamente ditos, podem ser vistos no gráfico abaixo:



Figura 14 – Dados SST e OD – ETE 1

Fonte: Do autor, 2022.

Outro exemplo pode ser visto, agora analisando-se a ETE 2. Uma forte correlação (0,7) pode ser enxergada entre o SST do licor misto e os mantos de lodo dos DSs com um atraso de 12 horas:

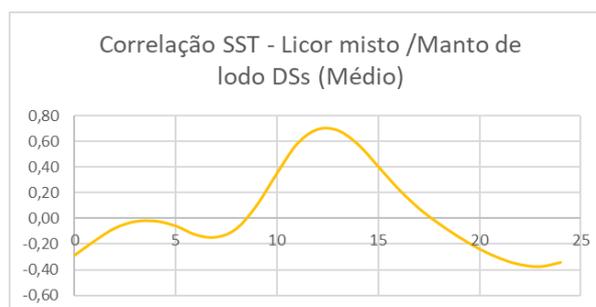


Figura 15 – Correlação do SST do licor misto com o Manto de lodo dos DSs

Fonte: Do autor, 2022.

## 5 Conclusões/Considerações finais e contribuições

Em paralelo ao trabalho de enxergar possíveis correlações e atuações no controle, também foi levantado junto a operação, os limites e set-points dos parâmetros analisados a fim de nos auxiliarem nas etapas futuras do trabalho. Segue um primeiro esboço para os limites da ETE 2:

Tabela 3 – Limites ou set-points desejados

TAG	Parâmetro	Local	Limites ou set-point desejado?
P4-AIT-101	pH	Esgoto desarenado	6,5 a 7,5
P4-AIT-102	Condutividade	Esgoto desarenado	450 a 1.600 µS/cm
P4-AIT-103	SST (imersão)	Esgoto desarenado	120 a 280 mg/L
P4-AIT-104	DQO	Esgoto desarenado	400 a 500 mg/L
P4-AIT-105	SST (imersão)	Efluente Primário	100 a 180 mg/L
P4-AIT-106	DQO	Efluente Final	24 a 60 mg/L
P4-LIT-301	Manta de Lodo	Decantador Prim.	≤ 1,5 m
P4-LIT-302	Manta de Lodo	Decantador Prim.	≤ 1,5 m
P4-AIT-109	SST (imersão)	Licor misto	1.500 a 3.500 mg/L
P4-AIT-110	SST (imersão)	Lodo de Retorno	3.000 a 5.000 mg/L
P4-AIT-111	ORP	Reator aeróbio 1	+ 80 a + 180 mV
P4-AIT-112	ORP	Reator aeróbio 1	+ 30 a + 100 mV
P4-AIT-113	ORP	Reator aeróbio 1	+ 100 a + 300 mV

P4-AIT-114	ORP	Reator aeróbio 2	+ 80 a + 180 mV
P4-AIT-115	ORP	Reator aeróbio 2	+ 30 a + 100 mV
P4-AIT-116	ORP	Reator aeróbio 2	+ 100 a + 300 mV
P4-AIT-117	ORP	Reator aeróbio 3	+ 80 a + 180 mV
P4-AIT-118	ORP	Reator aeróbio 3	+ 30 a + 100 mV
P4-AIT-119	ORP	Reator aeróbio 3	+ 100 a + 300 mV
P4-AIT-120	ORP	Reator aeróbio 4	+ 80 a + 180 mV
P4-AIT-121	ORP	Reator aeróbio 4	+ 30 a + 100 mV
P4-AIT-122	ORP	Reator aeróbio 4	+ 100 a + 300 mV
P4-AIT-123	Oxigênio Dissolvido	Reator aeróbio 1	1 a 3 mg/L
P4-AIT-124	Oxigênio Dissolvido	Reator aeróbio 1	0,1 a 2 mg/L
P4-AIT-125	Oxigênio Dissolvido	Reator aeróbio 1	1 a 3 mg/L
P4-AIT-126	Oxigênio Dissolvido	Reator aeróbio 2	1 a 3 mg/L
P4-AIT-127	Oxigênio Dissolvido	Reator aeróbio 2	0,1 a 2 mg/L
P4-AIT-128	Oxigênio Dissolvido	Reator aeróbio 2	1 a 3 mg/L
P4-AIT-129	Oxigênio Dissolvido	Reator aeróbio 3	1 a 3 mg/L
P4-AIT-130	Oxigênio Dissolvido	Reator aeróbio 3	0,1 a 2 mg/L
P4-AIT-131	Oxigênio Dissolvido	Reator aeróbio 3	1 a 3 mg/L
P4-AIT-132	Oxigênio Dissolvido	Reator aeróbio 4	1 a 3 mg/L
P4-AIT-133	Oxigênio Dissolvido	Reator aeróbio 4	0,1 a 2 mg/L
P4-AIT-134	Oxigênio Dissolvido	Reator aeróbio 4	1 a 3 mg/L
P4-AIT-135	NH3	Efluente Primário	Até 30 mg/L
P4-AIT-136	NH3	Licor misto	Até 25 mg/L
P4-AIT-137	Turbidez	Efluente tratado	Até 10 NTU
P4-LIT-501	Manto de Lodo	Decantador DS1	≤ 50 cm
P4-LIT-502	Manto de Lodo	Decantador DS2	≤ 50 cm
P4-LIT-503	Manto de Lodo	Decantador DS3	≤ 50 cm
P4-LIT-504	Manto de Lodo	Decantador DS4	≤ 50 cm
P4-AIT-142	SST (imersão)	Lodo Aflu. Digestor	23.000 a 38.000 mg/L
P4-AIT-143	SST (imersão)	Lodo condicionado	23.000 a 35.000 mg/L
P4-AIT-144	SST (imersão)	Filtrado FPP 1 a 3	Até 220 mg/L
P4-AIT-145	NH3	Filtrado FPP 1 a 3	Até 350 mg/L
P4-AIT-146	Condutividade	Filtrado FPP 1 a 3	200 a 1.500 µS/cm
P4-AIT-147	Cloro (inserção)	ETA Reuso	>1 e <10 mg/L
P4-AIT-148	Turbidez	ETA Reuso	Até 10 NTU

Fonte: O autor, 2022

Muitas outras correlações podem ser obtidas com suas respectivas diferenças horárias, iniciando assim um banco de dados de como operar de maneira otimizada as Estações de Tratamento, prevendo falhas e antecipando atuações dos controles.

Etapas futuras do trabalho, devem fazer essa análise mostrada e outras de forma automática, haja vista que esses dados estarão chegando aos controladores, sendo enviados ao supervisório e armazenados no banco de dados das plantas.

A ideia nesse ponto do trabalho é utilizar o software da empresa BirminD que possui soluções para uso das tecnologias habilitadoras da Indústria 4.0. Espera-se utilizar dos módulos: “Relevância”, “Otimização”, “Simulação” e “Detecção de anomalias”. Abaixo podemos visualizar o módulo “simulação” no add-in para Excel utilizando o mesmo banco de dados que foi usado para obter as correlações observadas anteriormente:

Tabela 4 – Simulação de relevância add-in para Excel BirMind em dados analíticos do tratamento de esgoto

output	pH - Esgoto bruto	
inputs_relevance	SST (imersão) - Esgoto desarenado	13,05%
	ORP - Reator Aeróbio 1A	12,38%
	ORP - Reator Aeróbio 1B	10,41%
	ORP - Reator Aeróbio 2A	9,27%
	ORP - Reator Aeróbio 1B	8,65%

Oxigênio Dissolvido - Reator Aeróbio 1B	6,87%
Oxigênio Dissolvido - Reator Aeróbio 2A	6,63%
Oxigênio Dissolvido - Reator Aeróbio 2B	5,37%
Oxigênio Dissolvido - Reator Aeróbio 1A	4,89%
Oxigênio Dissolvido - Reator Aeróbio 2B	4,63%
Oxigênio Dissolvido - Reator Aeróbio 1B	3,63%
SST (imersão) - Filtrado FPP 1 a 3	3,04%
Manto de Lodo - Tanque DS5	2,59%
Manto de Lodo - Tanque DS6	2,51%
Manto de Lodo - Tanque DS1	2,11%
Manto de Lodo - Tanque DS3	2,06%
Manto de Lodo - Tanque DS2	1,92%

Fonte: O autor, 2022

Observado o resultado da análise, temos que diversos dados analíticos possuem correlação com o pH nas porcentagens que foram apresentadas, ou seja, é possível entender quais dados possuem mais correlação em relação aos outros.

Outro módulo interessante é de “Otimização” que determina como devem ser os outros parâmetros, afim de obter um set-point específico em determinada análise:

Tabela 5 - Simulação de otimização add-in para Excel BirMind em dados analíticos do tratamento de esgoto

output results	P1-AIT-101	
	P1-AIT-603	3,35
	P1-LIT-439	1,04
	P1-LIT-438	1,04
	P1-LIT-436	1,31
	P1-LIT-435	0,5
	P1-LIT-434	0,92
	P1-AIT-421	5,43
	P1-AIT-419	1
	P1-AIT-418	5,27
	P1-AIT-417	4,86
	P1-AIT-415	0,66
	P1-AIT-414	4,8
	P1+AIT+411	108,94
	P1-AIT-410	282,16
	P1-AIT-409	245,43
	P1-AIT-408	234,09
	P1-AIT-103	182,94
entries_score	7	
optimal_point	7	

Fonte: O autor, 2022

Nesse passo também foi possível entender e projetar a dimensão dos dados analíticos dentro da Unidade, visando saber de necessidade de novas alternativas ou não para o armazenamento dos dados existentes hoje, ou seja, se podemos considerar como Big Data os dados analisados:

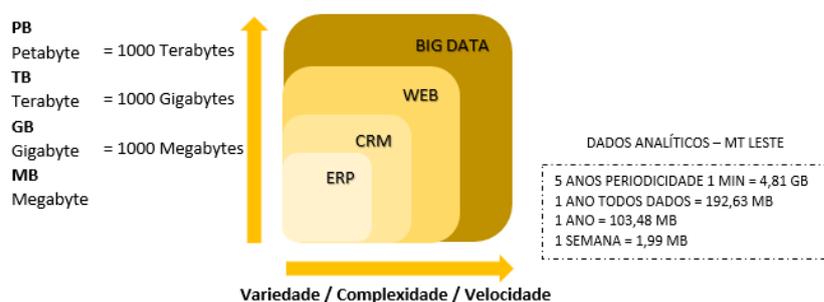


Figura 16 – Projeção de dados analíticos  
 Fonte: Paulucci & Rubim, 2022

Outro conceito importante é entender onde a companhia se encontra e onde pode se posicionar no médio prazo com o uso do Data Science, passando do Business Intelligence (BI) para o Business Analytics, com análises preditivas e prescritivas:

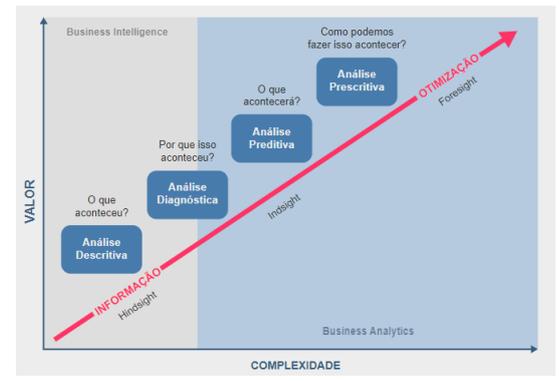


Figura 17 – Business Intelligence (BI) para o Business Analytics  
 Fonte: Paulucci & Rubim, 2022

Para isso será fundamental a intersecção de algumas áreas do conhecimento afim de se chegar a esse Data Science, são essas: Matemática e Estatística, Ciência da Computação e o Domínio do conhecimento.

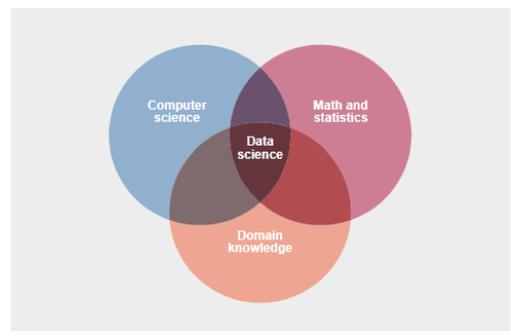
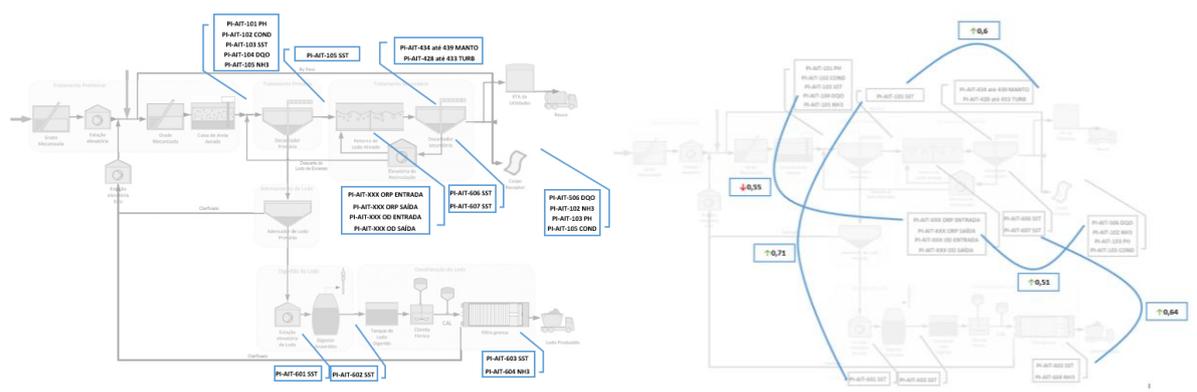


Figura 18 – Tripé do Data Science  
 Fonte: Paulucci & Rubim, 2022

Por fim a ideia é que possamos, em etapas futuras do trabalho, desenhar algo como um mapa das Estações de Tratamento de Esgoto, com base nos dados analíticos agora analisados, e elencar as correlações que foram enxergadas como os respectivos “delays”:



Figuras 19 e 20 – Fluxograma do processo com localização dos dados analisados, elencando correlações (exemplo)

Fonte: Do autor, 2022

## 6 Referências

BBC. Eles pararam de reclamar dos governos e estão usando o celular para melhorar a política. Disponível em: <https://www.bbc.com/portuguese/brasil-36660936>. Acesso em: 24 mar. 2022.

CHOI, S. S.; KANG, G.; JUNG, K.; KULVATUNYOU, B. & MORRIS, K.C. Applications of the factory design and improvement reference activity model. In: INTERNATIONAL CONFERENCE ON ADVANCES IN PRODUCTION MANAGEMENT SYSTEMS (APMS), 2016. Iguassu Falls, Brazil. Disponível em: <https://hal.inria.fr/hal-01615741/document>. Acesso em: 14 mar. 2022.

BRASIL. Presidência da República. Lei n. 14.026, de 15 de julho de 2020. Atualiza o marco legal do saneamento básico. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2019-2022/2020/lei/114026.htm](http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/lei/114026.htm). Acesso em: 14 mar. 2022.

SÃO PAULO (ESTADO). Anuário de energéticos por municípios do Estado de São Paulo: 2021 ano base 2020. Disponível em: [http://dadosenergeticos.energia.sp.gov.br/portalicev2/intranet/BiblioVirtual/diversos/anuario\\_energetico\\_municipio.pdf](http://dadosenergeticos.energia.sp.gov.br/portalicev2/intranet/BiblioVirtual/diversos/anuario_energetico_municipio.pdf). Acesso em: 07 mai. 2018.

DRATH, R.; HORCH, A. Industrie 4.0: Hit or hype? IEEE industrial electronics magazine, v. 8, n. 2, p. 56–58, 2014.

PAULUCCI, A.; RUBIM, L. FIAP - Nano courses. Big Data & Analytics. Disponível em: <https://on.fiap.com.br/local/nanocourses/index.php> Acesso em: 27 jun. 2022

FIRJAN. Indústria 4.0: panorama da Inovação. 2016. Disponível em: <https://www.firjan.com.br/lumis/portal/file/fileDownload.jsp?fileId=2C908A8A555B47FF01557D8802C639A4>. Acesso em: 09 mar. 2022.

JORDÃO, E. P.; PESSOA, C. A. Tratamento de esgotos domésticos. 7. ed. Rio de Janeiro: ABES, 2014.

GARTNER. Tecnologia da Informação. 2022. Disponível em: <https://www.gartner.com/en/information-technology/glossary/big-data>. Acesso em: 14 mar. 2022.

GOEKING, W., Da máquina a vapor aos softwares de automação, Revista O Setor Elétrico, Edição 52, Brasil, maio de 2010. Disponível em [https://www.voltimum.com.br/sites/www.voltimum.com.br/files/memoria\\_maio\\_10.pdf](https://www.voltimum.com.br/sites/www.voltimum.com.br/files/memoria_maio_10.pdf). Acesso em: 10 mar. 2022.

GOMES, Dennis S. Inteligência Artificial: conceitos e aplicações. Revista Olhar Científico. Faculdades Associadas de Ariquemes, v. 01, n.2, ago./dez. 2010. Disponível em: <https://www.academia.edu/download/48312264/49-148-1-PB.pdf>. Acesso em: 10 mar. 2022.

GOOGLE TRENDS. Big data. Disponível em: <https://trends.google.com.br/trends/explore?date=all&geo=BR&q=BIG%20DATA>. Acesso em: 08 mar. 2022.

KAGERMANN, H.; WAHLSTER, W. & HELBIG, J. Recommendations for implementing the strategic initiative industrie 4.0: final report of the Industrie 4.0, 2013. Disponível em: <https://www.din.de/blob/76902/e8cac883f42bf28536e7e8165993f1fd/recommendations-for-implementing-industry-4-0-data.pdf>. Acesso em: 14 mar. 2022.

METCALF & EDDY. Wastewater engineering: treatment and reuse. 4. ed. [s. l.]: McGraw-Hill, 2003.

MITCHELL, Tom. Machine learning. New York: McGraw-Hill, 1997.

NUNES, Guilherme B. Análise de maturidade 4.0 do núcleo de tecnologias tridimensionais do CTI Renato Archer. Disponível em: [https://www.gov.br/cti/pt-br/publicacoes/producao-cientifica/seminario-pci/x\\_seminario\\_pci-2020/pdf/seminario-2020\\_paper\\_25.pdf](https://www.gov.br/cti/pt-br/publicacoes/producao-cientifica/seminario-pci/x_seminario_pci-2020/pdf/seminario-2020_paper_25.pdf). Acesso em: 27 jun. 2022.

ALTUS. Curso de introdução a automação. Disponível em: <https://www.altus.com.br/post/100/curso-de-introducao-a-automacao--5baula-01-5d>. Acesso em: 23 jun. 2022.

SACOMANO, J.B. et al. Indústria 4.0: conceitos e fundamentos. São Paulo: Blucher, 2018.

SAMUEL, Arthur L. Some studies in machine learning using the game of checkers. IBM Journal of Research and Development, v. 3, n. 3, pp. 210-229, July, 1959. DOI: 10.1147/rd.33.0210. Disponível em: <https://ieeexplore.ieee.org/document/5392560>. Acesso em: 23 jun. 2022.

SCHUMACHER, A.; EROL, S.; SIHN, W. A maturity model for assessing Industry 4.0 readiness and maturity of manufacturing enterprises. Procedia CIRP, v. 52, p. 161-166, 2016. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2212827116307909>. Acesso em: 14 mar. 2022.

STAIR, Ralph M.; REYNOLDS, George W. Princípios de sistemas de informação. 9. ed. São Paulo: Cengage Learning, 2013.

VAN HAANDEL, A. C.; MARAIS, G. V. R. O Comportamento do Sistema de Lodo Ativado. Campina Grande: Epgraf, 1999.