

# Trabalho de Sistemas de Apoio à Decisão Business Intelligence

Grupo Les Amis du Codage:  
Celso Antonio Uliana Junior, Felipe Salles Lopes, Lucas Avanzi

RGA: 2014.1907.003-7, RGA: 2016.1907.032-4, RGA: 2016.1907.024-3

## 1. Relatório da Quarta Entrega

### 1.1. Introdução

Este relatório tem como objetivo apresentar a execução da tarefa da quarta entrega parcial proposta no cronograma de entregas, além de seu status e os novos elementos adicionados, alterados e corrigidos no projeto.

### 1.2. Status

As tarefa *“Implementação de algoritmos de aprendizado de máquinas para séries temporais e seus resultados em Python utilizando a biblioteca Scikit-learn”* foram inicializadas, porém não finalizadas. Os responsáveis por estas tarefas foram os integrantes Celso A. Uliana e Felipe S. Lopes.

### 1.3. Execução

A tarefa, antes de ter sido iniciada, foi feito uma preparação para esta. Como explicado na reunião anterior com o professor, trocamos o objetivo 2 no qual constava que iríamos prever os próximos números de morte pela doença Covid19 a partir da comparação da análise anterior feita no primeiro objetivo com base no número de leitos disponíveis no Brasil. Por conta da complexidade da correlação entre o número de leitos disponíveis e o número de morte e, também, por conta da falta de confiabilidade nessa primeira fonte de dados dos leitos, foi mudado o segundo objetivo para realizar a predição do número de casos de contágio com base no índice de isolamento social no país.

Para tal buscamos os dados pelo fornecedor direto da fonte: [inloco](#). No entanto, o site não nos disponibiliza os dados de forma fácil e, por conta disso, foi preciso realizar a obtenção dos dados de forma manual, o que deixou a tarefa um pouco mais trabalhosa. Além disso, o site apenas mostra de forma diária os dados em um intervalo de apenas uma semana.

Após a obtenção dos dados, criamos uma nova tabela no nosso DataWarehouse para acomodar esses dados e, em consequência, foi necessário alterar a modelagem do DW.

Adicionamos os códigos SQL feitos para essa tarefa junto aos códigos anteriores e injetamos os dados no DW. Ademais foi atualizado a quantidade de dias na tabela *tempo* no DW.

Na parte de aprendizado de máquina, utilizamos a biblioteca Scikit-learn e tomamos como base o código mostrado nas aulas para fazer um pequeno pedaço do segundo objetivo. Porém, como os dados de isolamento social é muito pequeno por enquanto (pelo motivo descrito anteriormente) e temos uma dúvida na qual enviaremos um e-mail logo em seguida, não conseguimos o resultado esperado, porém o código está encaminhado para o nosso objetivo e já produz um resultado (não esperado). Os códigos SQL e Python dessa tarefa já está disponível no nosso repositório na branch [develop](#).

	id [PK] integer	data_reg integer	taxa_isolamento numeric
1	7	166	48.38
2	8	167	39.33
3	9	168	39.56
4	10	169	37.75
5	11	170	39.01
6	12	171	35.02
7	13	172	39.37

**Figura 1. Tabela indice\_isolamento\_social**

É previsto que finalizaremos essa parte nessa semana ainda e entregaremos já na próxima entrega prevista.

## 1.4. Alterações

Houveram algumas alterações partes do projeto. Como sempre, tais alterações serão descritas no documento CHANGELOG.md que será enviado em anexo junto a este relatório. Uma dessas entregas vale a pena ser mencionada.

### 1.4.1. Modelagem do Datawarehouse

