

Case - Calina

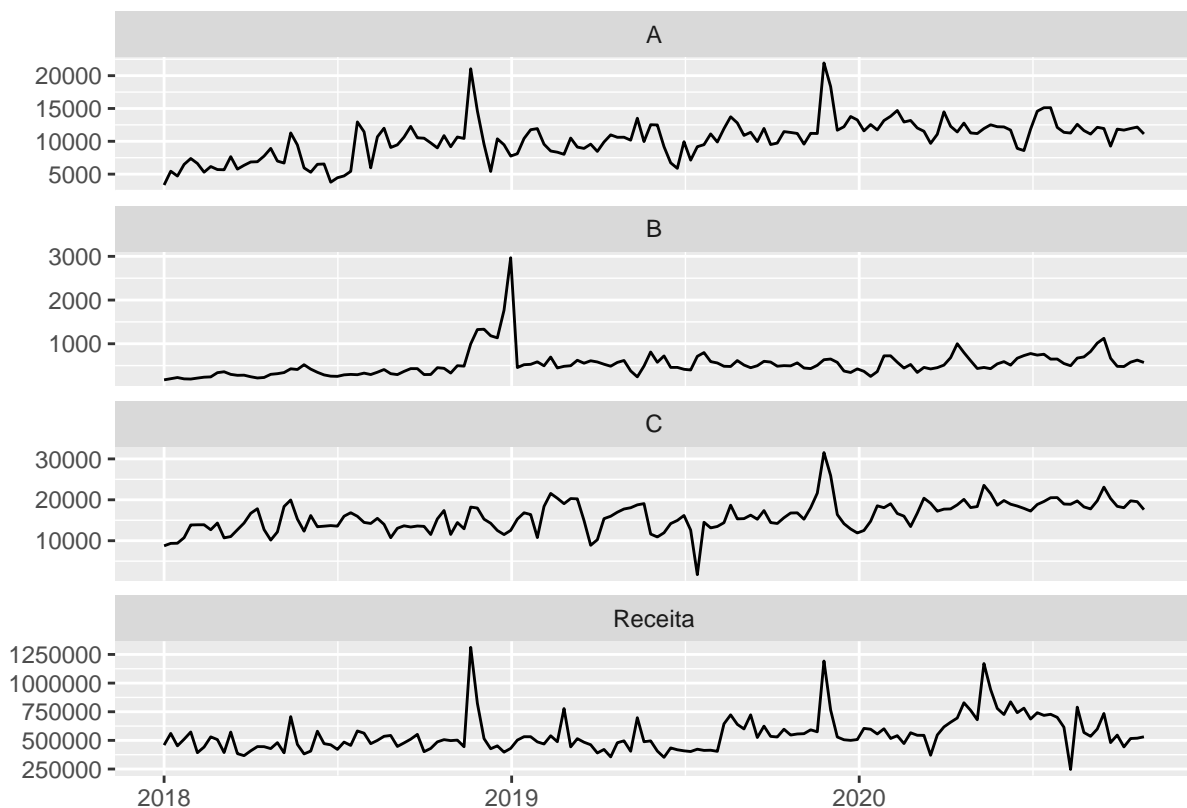
Roger Gregório Marcondes

Instruções do case:

Em uma agência de Marketing Digital uma das épocas mais importante para o ramo de ecommerce é a “Black Friday”, período sazonal em que muitos dos clientes se planejam com promoções e ações através das mídias pagas para chamar a atenção dos usuários. Pensando nessa data muito especial, um cliente da Calina solicitou uma análise para prevermos qual será a receita da Black Friday de 2020.

O banco de dados enviado contém dados de 3 mídias em que o cliente investe (Mídia A, B e C) e o total da receita gerada no site por semana, desde a primeira semana de 2018 até a última semana de outubro de 2020.

Para responder ao cliente análise o banco de dados, crie um modelo teste e um modelo final que deve prever as próximas 4 semanas, respectivas ao mês de novembro de 2020 (a última semana é a semana da Black Friday).



A figura mostra as variáveis que podem ser úteis para prever as próximas 4 semanas da Black Friday, respectivas ao mês de novembro de 2020. Estas são variações percentuais semanais da receita dos anos de 2018, 2019 e 2020, e variações semanais das ações através das mídias pagas

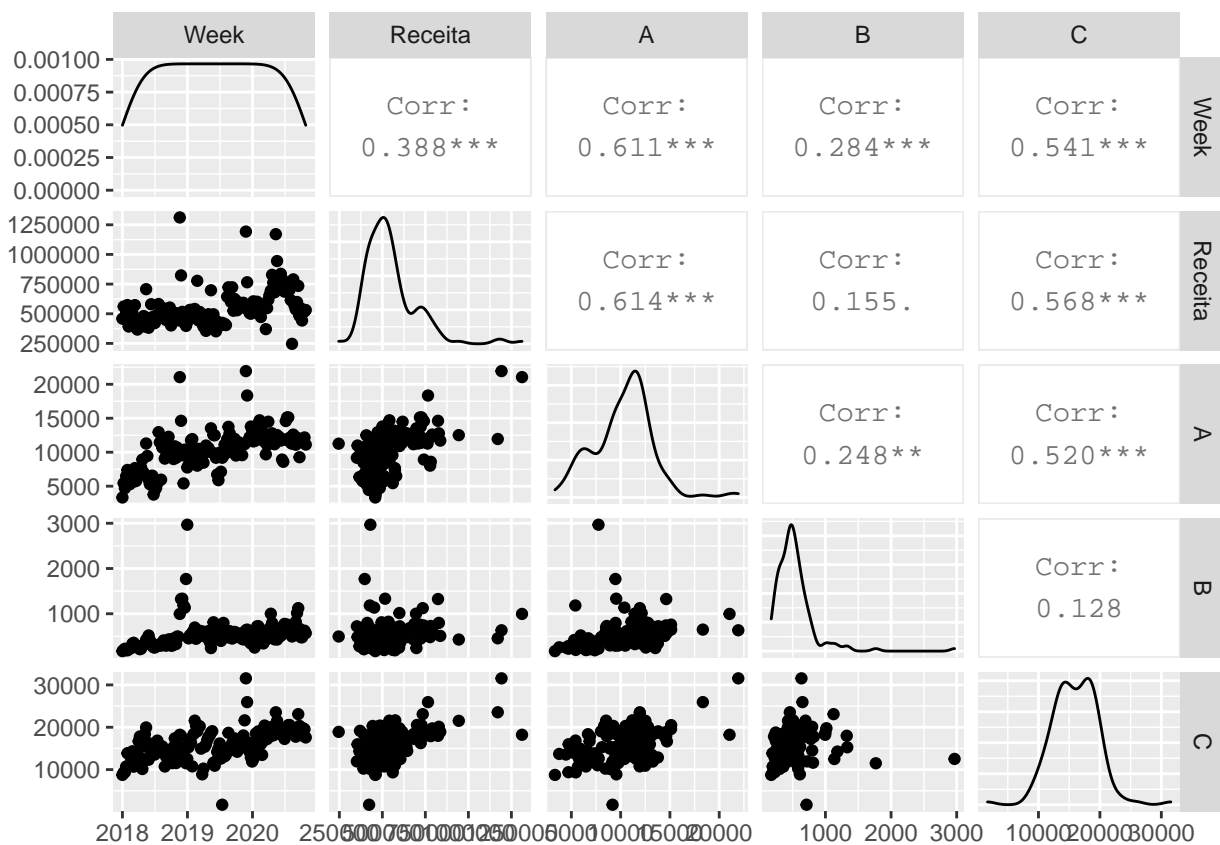
(A, B e C), que tem como objetivo chamar a atenção dos usuários no período sazonal em que muitos clientes se planejam com as promoções oferecidas.

Quando analisamos a média A podemos notar uma tendência crescente, presença de ciclos e sazonalidade presente nos períodos de Black Friday para os anos de 2018 e 2019, para a média B temos uma tendência constante, presença de ciclos e sazonalidade presente somente no período de Black Friday para o ano de 2018 e para a média C temos uma tendência crescente, presença de ciclos e sazonalidade presente somente no período de Black Friday para o ano de 2019.

Verifica-se os períodos correspondentes a Black Friday de 2018 e 2019 (4 semanas do mês de novembro), para a variável resposta Receita.

Table 1: Receita da Black Friday em 2018 e 2019

	2018	2019
1ªSemana	502405.2	558475.9
2ªSemana	444800.3	592416.8
3ªSemana	1311737.6	574789.8
4ªSemana	823460.9	1192699.9



A figura é uma matriz do gráfico de dispersão de três variáveis. A segunda coluna mostra os relacionamentos entre a variável de previsão (Receita) e cada um dos preditores (Médias). Os gráficos de dispersão mostram relações positivas. A força dessas relações é mostrada pelos coeficientes de correlação na segunda linha. Os demais gráficos de dispersão e coeficientes de correlação mostram as relações entre os preditores.

A saída fornece informações sobre o modelo ajustado inicial.

##

```
## Call:
## tslm(formula = Receita ~ A + B + C, data = DadoST)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -367969  -63796  -18133   46282   485200
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 88712.176  44806.754   1.980  0.0496 *
## A             22.604     3.799   5.950 1.95e-08 ***
## B              1.486    31.267   0.048  0.9622
## C             14.300     3.006   4.757 4.72e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 114800 on 144 degrees of freedom
## Multiple R-squared:  0.4614, Adjusted R-squared:  0.4502
## F-statistic: 41.13 on 3 and 144 DF,  p-value: < 2.2e-16
```

Então a equação do modelo ajustado inicial em função do tempo dado em semanas será:

$$y_t = 88712.176 + 22.604A_t + 1.486B_t + 14.300C_t + \varepsilon_t$$

Mas como a média B não é significativa para o modelo. Verificamos outras possibilidades:

Table 2: Escolha do melhor modelo de previsão

	BIC	A	B	C
5	3464.666	1	0	1
7	3469.661	1	1	1
4	3481.271	1	0	0
6	3486.267	1	1	0
1	3493.703	0	0	1
3	3497.194	0	1	1
2	3547.658	0	1	0

Sendo assim escolhemos o modelo “5” com $BIC = 3464.666$ que contém as médias A e C.

Ganho do modelo final (em função das médias A e C) em relação ao modelo inicial (em função das médias A, B e C) é de:

```
## [1] 1.001442
```

A saída fornece informações sobre o modelo ajustado final.

```
##
## Call:
## tslm(formula = Receita ~ A + C, data = DadoST)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
```

```
## -368061 -64046 -18030 46863 485470
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 89109.839  43867.076   2.031   0.044 *
## A           22.642     3.698    6.123 8.18e-09 ***
## C           14.300     2.996    4.774 4.38e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 114400 on 145 degrees of freedom
## Multiple R-squared:  0.4614, Adjusted R-squared:  0.454
## F-statistic: 62.12 on 2 and 145 DF,  p-value: < 2.2e-16
```

Sendo assim a equação do modelo ajustado final em função do tempo dado em semanas será:

$$y_t = 89109.839 + 22.642A_t + 14.300C_t + \varepsilon_t$$

Com o modelo pré estabelecido, temos que a intenção de investimento não foi informado para as mídias (A, B e C) para as 4 semanas de novembro de 2020, verifica-se a possibilidade de injeção dos investimentos nas mídias A_t e C_t em função do tempo de duas formas, na primeira consideramos que os investimentos nas mídias A e C do ano 2019 serão mantidos e injetados para a previsão da receita do modelo ajustado final, na segunda é fazer uso das características de tendência e sazonalidade do histórico de investimentos de toda a base de dados disponibilizada.

Neste caso, considera as variáveis falsas correspondes a todas as semanas do ano na forma de regressão linear (tomando como exemplo a mídia A):

$$A_t = A_0 + C_0Semana1_t + \dots + C_{52}Semana52_t + \varepsilon_t$$

Pelos modelos de Regressão Linear, podemos estimar o investimento para as 4 semanas de novembro de 2020 com base nas informações de tendência e sazonalidade através do histórico de investimentos nas mídias (A, B e C).

Table 3: Estimativa dos investimentos em novembro de 2020

A	B	C
13414.81	639.643	18570.34
14125.13	627.633	19206.77
19421.91	919.533	23647.86
21583.64	1147.913	28487.52

Então obtemos as projeções da receita considerando os dois cenários de investimentos descritos acima e seus respectivos gráficos de tendência.

Previsão da Receita para o modelo ajustado final
 dado pelas médias A, B e C projetadas

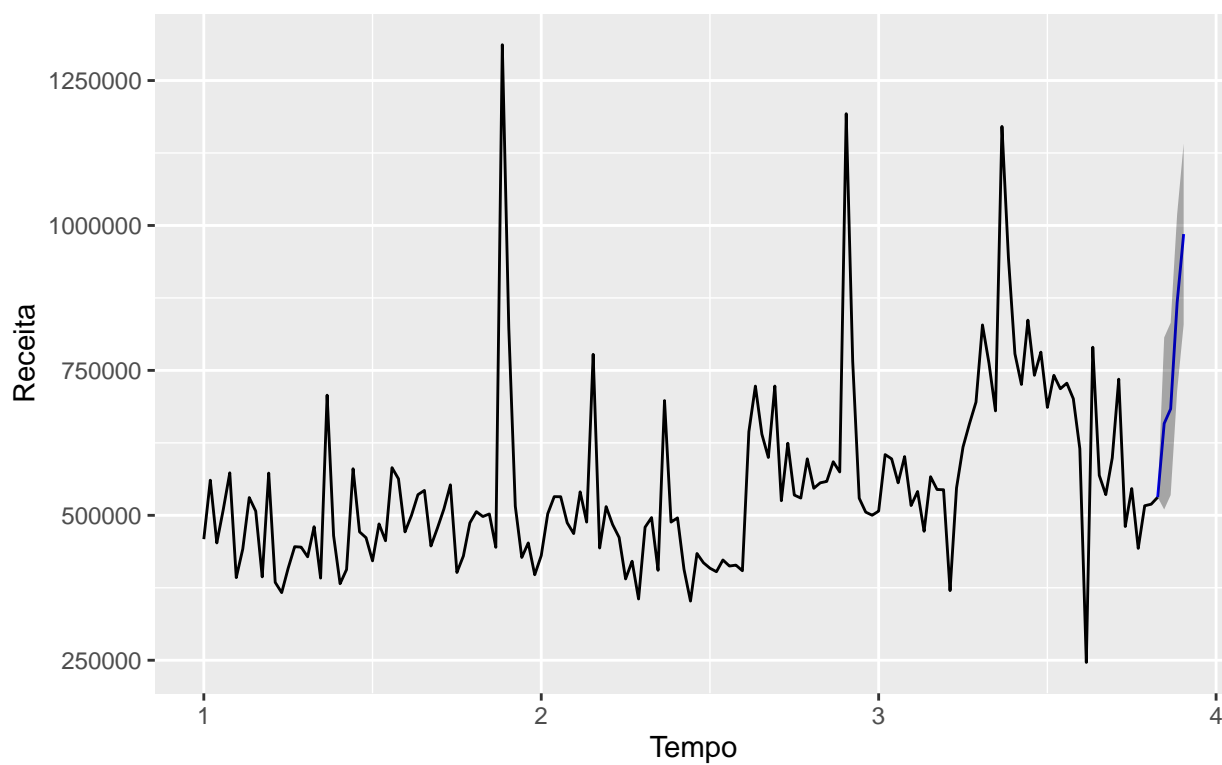
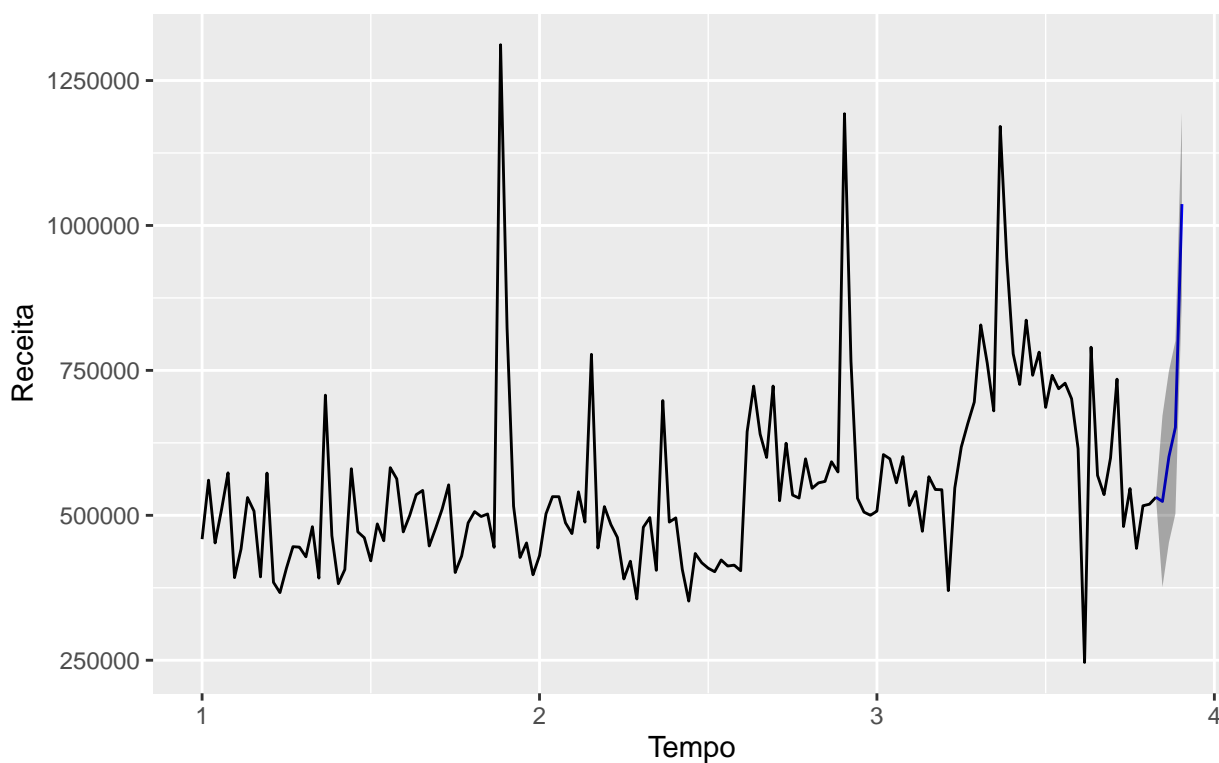


Table 4: Projeção de receita para a Black Friday de novembro de 2020

	Receita em 2019	Receita estimada em 2020
1ªSemana	523863.2	658404.3
2ªSemana	601126.4	683588.6
3ªSemana	651726.8	867026.5
4ªSemana	1036895.0	985179.2

Previsão da Receita para o modelo ajustado final
em relação aos investimentos de 2019



Como todo o mercado sofreu e vem sofrendo os impactos do coronavírus (Covid-19), podemos notar uma diminuição da receita em relação a primeira semana de Black Friday de 2020 para ambas projeções.