

Challenge - IntelliDeere - Disciplina: Statistical Computing with R - 1TIAR 2023:

- Arthur Coutinho Santos --- RM: 97804
- Camilly Alves --- RM: 550210
- Guilherme Garcia Paschoalinoto --- RM: 99221
- João Vitor de Andrade Martins --- RM: 98744
- Murilo Krauss --- RM: 98262

2E) Interpretar todos os resultados, destacando os principais insights.

Possuímos 2 dataset's um de Pré-Venda, que possui 8 colunas e 104 registros, e outro de Pós-Venda, que possui 114 registros e 8 colunas.

1. Pré-Venda:

a. Entendendo o dataset:

O nosso dataset de pré-venda é constituído de 8 colunas, que são:

- Data da venda
- Localização
- Área Plantada
- Tipo de cultura
- Preço médio
- Método de pagamento
- Número de vendas
- Condição dos veículos

E cada linha do dataset se refere a uma venda realizada de equipamento.

b. Análise das correlações:

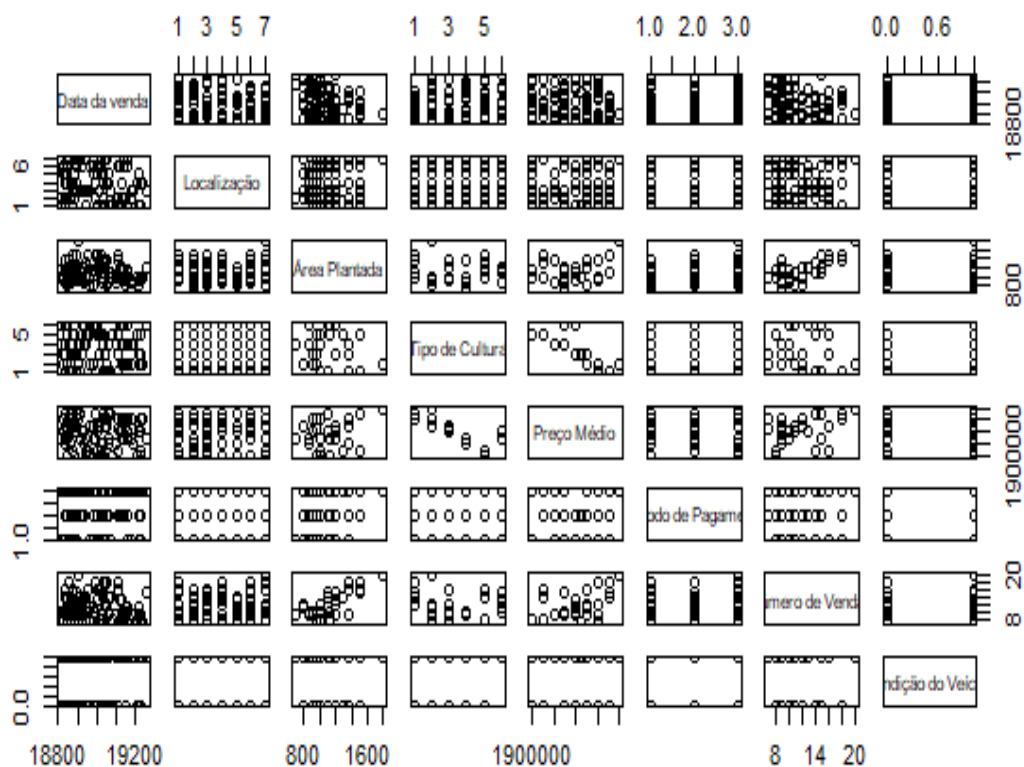
	Data da venda	Localização	Área Plantada	Tipo de Cultura	Preço Médio	Método de Pagamento	Número de Vendas	Condição do Veículo
Data da venda	1	-0.13757013	-0.22286358	0.01206286	-0.07410599	0.12444118	-0.21664155	0.14347413
Localização	-0.13757013	1	0.02154688	-0.02278845	0.07690674	0.06037485	0.09963804	-0.02367100
Área Plantada	-0.22286358	0.02154688	1	0.00304104	0.16764003	0.06596902	0.78529946	-0.00505525
Tipo de Cultura	0.01206286	-0.02278845	0.00304104	1	0.76718354	0.01787693	-0.25874155	0.38975448
Preço Médio	-0.07410599	0.07690674	0.16764003	0.76718354	1	0.12993636	0.38810931	-0.08103904
Método de Pagamento	0.12444118	0.06037485	0.06596902	0.01787693	0.12993636	1	-0.04403574	0.02863789
Número de Vendas	-0.21664155	0.09963804	0.78529946	0.25874155	0.38810931	-0.04403574	1	-0.20160489
Condição do Veículo	0.14347413	-0.02367100	-0.00505525	0.38975448	-0.08103904	0.02863789	-0.20160489	1

Com base nos valores fornecidos, percebemos que a data de venda não possui uma correlação forte com nenhuma variável, no indicando que a data da venda não possui uma relação direta com as variáveis, o mesmo se encaixa para a variável localização, pois ela apresentou relações fracas também.

Analisamos que, uma área de plantação tem uma correlação com os preços médios mais altos e número de vendas maiores mais elevados, ou seja podemos imaginar que por ser

um terreno maior o agricultor busca maquinas mais “potentes” e um numero maior de maquinas, para poder cobrir todo o seu terreno.

Em relação ao método de pagamento notamos que dependendo do método do pagamento ele possui um preço menor, variando de caso a caso, mas imaginamos também que se encaixa em caso como um valor mais elevado podem necessitar de um financiamento ou coisa do tipo.



2. Pós-Venda:

a. Entendendo o dataset:

O nosso dataset de pós-venda é constituído de 8 colunas, que são:

- Data da compra
- Localização
- Equipamento
- Tipo de serviço
- Tempo de Atendimento (horas)
- Peças substituídas
- Custo
- Cliente

b. Análise das correlações:

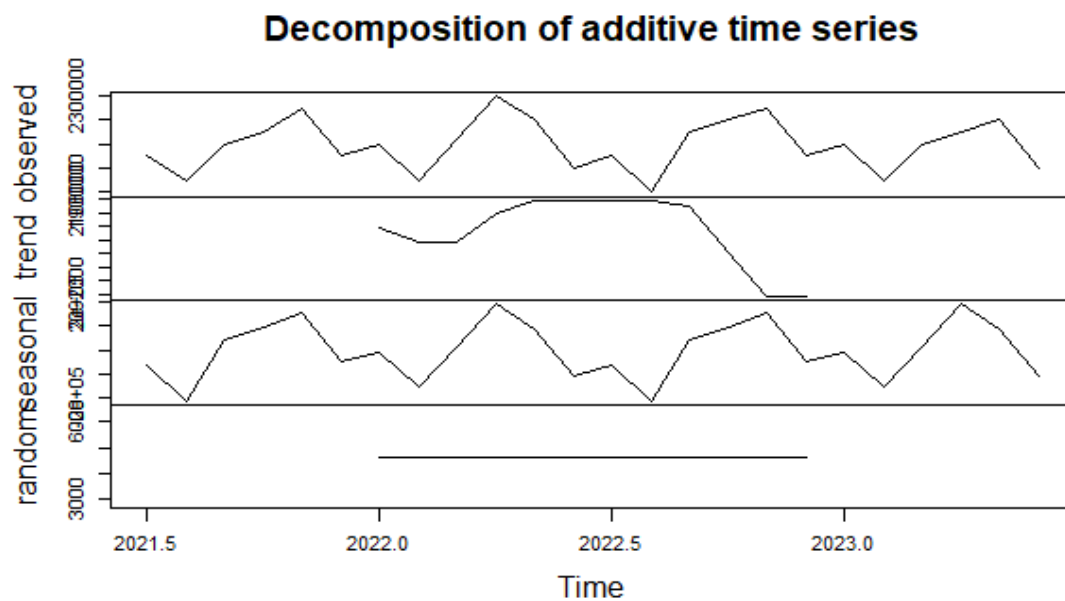
	Data da compra	Localização	Equipamento	Tipo de Serviço	Tempo de Atendimento	Peças Substituídas	Custo	Cliente
--	----------------	-------------	-------------	-----------------	----------------------	--------------------	-------	---------

Data da compra	1	0.02544368	-0.13886752	0.03545946	0.05231637	0.03810543	0.04895699	-0.05042060
Localização	0.02544368	1	0.05566962	-0.15218146	0.06517564	-0.02741604	0.07594189	-0.06021882
Equipamento	-0.13886752	0.05566962	1	-0.04071677	0.09831967	0.04165820	0.07682761	-0.05789195
Tipo de Serviço	0.03545946	-0.15218146	-0.04071677	1	-0.74279059	0.33126711	-0.24893020	0.04638587
Tempo de Atendimento	0.05231637	0.06517564	0.09831967	-0.74279059	1	0.25544815	0.71639617	0.03593503
Peças Substituídas	0.03810543	-0.02741604	0.04165820	0.33126711	0.25544815	1	0.58125003	0.03050701
Custo	0.04895699	0.07594189	0.07682761	-0.24893020	0.71639617	0.58125003	1	0.14694237
Cliente	-0.05042060	-0.06021882	-0.05789195	0.04638587	0.03593503	0.03050701	0.14694237	1

Com base na análise dos valores da tabela de correlação linear, notamos que 2 variáveis não têm influência significativa nos valores das nas outras variáveis, são elas data da compra e localização.

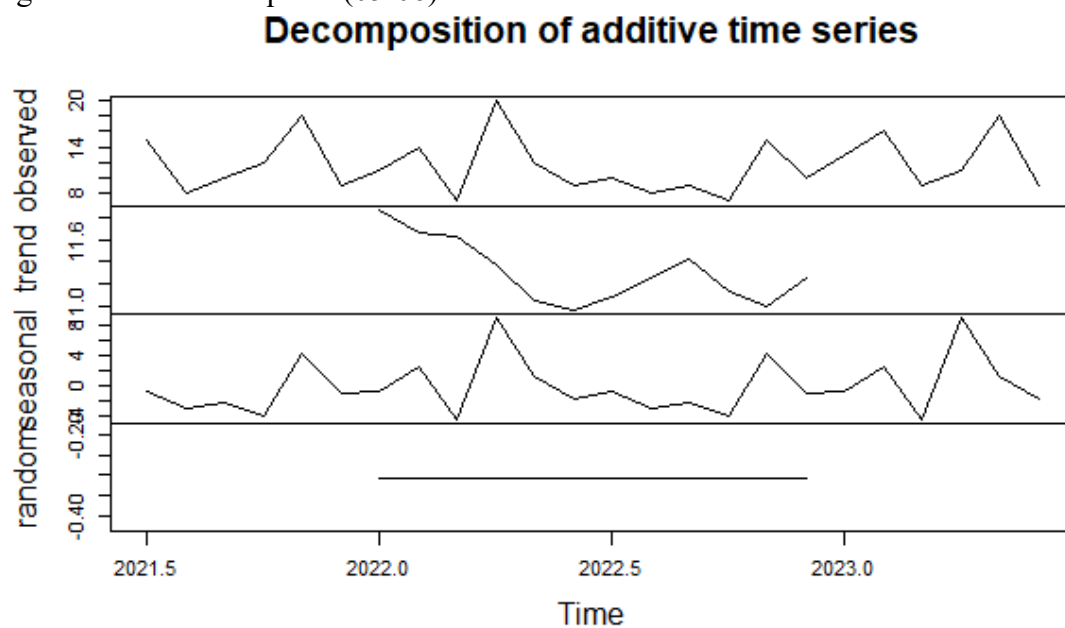
Percebemos também que a variável que mais interfere nas outras é o tipo de serviço feito na manutenção do equipamento, tendo cada serviço um tempo de atendimento (TMA) e um custo, mas também uma correlação positiva de que vai ser usado peças, quer dizer que nas manutenções vai ser utilizados peças novas ou algo do tipo.

Há também uma relação direta entre TMA e Custo, ou seja, de acordo com o TMA pode ser que o valor da manutenção varie, assim como não tem como pré definir um TMA e um custo de manutenção pra cada equipamento, precisa saber qual manutenção é necessária.



c. Número de Vendas

Ao longo do 1 semestre de 2022 tivemos uma grande queda no numero de vendas, podem ser por fatores extras, e temos uma sazonalidade muito forte e constante, notada por uma queda no inicio do ano (02-04) acompanhada de uma grande subida (03-05) seguida de uma leve queda (05-06)

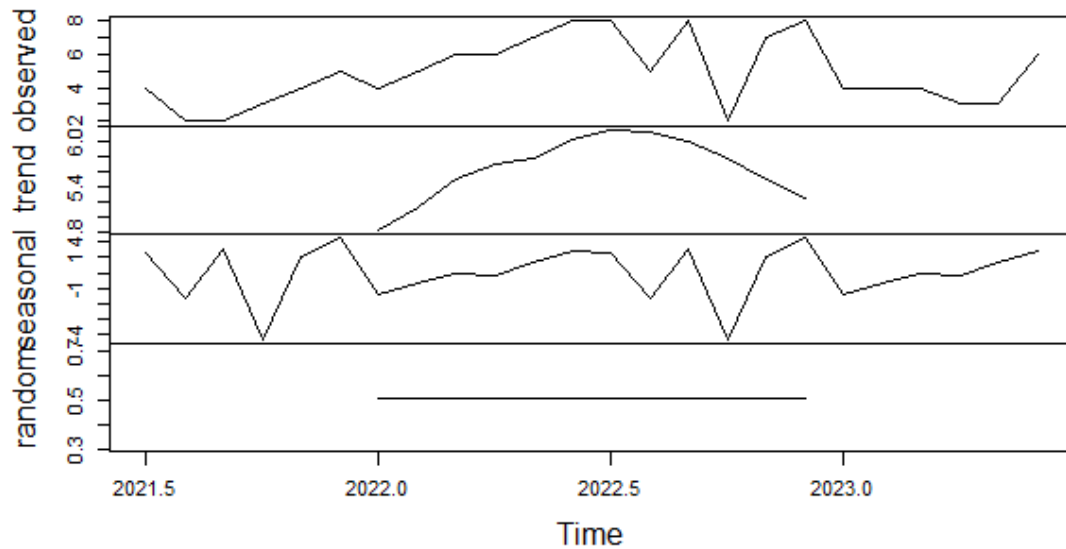


2. Pós-Vendas

a. Tempo de Atendimento (em horas)

Em relação ao tempo de atendimento ele teve uma tendência de alta no meio de 2022, Porém podemos notar uma sazonalidade, no final do ano sempre há uma queda e uma alta na busca por manutenções, provavelmente deve ser um período entre safras ou pré-safras

Decomposition of additive time series



b. Custo

Olhando os gráficos de custo podemos notar uma tendência de subida assim como o TMA e depois uma grande queda, já sazonalidade se converge quase na mesma linha do TMA, pois uma manutenção deve ser paga. Com um ruído constante, isso indica que não há padrões ou estrutura significativa não explicada pela tendência e sazonalidade

Decomposition of additive time series

