

## ***Análise de Regressão Linear Simples para Precisão de Vendas Online***

João Capelas, Miguel Velez, Ricardo Sousa

2ºL\_EI-IG-01, 201901812

2ºL\_EI-IG-01, 201900678

2ºL\_EI-IG-01, 201901874

**Resumo:** O trabalho irá retratar a análise feita a um conjunto de dados, sendo esse conjunto a relação entre o investimento em publicidade online e as vendas geradas online. O relatório do trabalho consiste numa análise descritiva sobre o conjunto de dados, recorrendo à regressão linear simples.

### **Palavras-chave:**

1. Análise de Dados
2. Regressão Linear
3. Resíduos
4. Publicidade
5. Vendas

## Índice

<b>Índice de Figuras .....</b>	<b>2</b>
<b>Índice de Tabelas .....</b>	<b>2</b>
<b>1. Introdução.....</b>	<b>3</b>
<b>2. Dados.....</b>	<b>3</b>
<b>3. Análise das variáveis.....</b>	<b>3</b>
<b>4. Conclusões.....</b>	<b>5</b>
<b>5. Referências bibliográficas.....</b>	<b>5</b>

## Índice de Figuras

<b>Figura 1 - Diagrama de dispersão das duas variáveis em análise .....</b>	<b>3</b>
<b>Figura 2 - Diagrama de dispersão das duas variáveis em análise com reta regressão .....</b>	<b>4</b>
<b>Figura 3 - Diagrama de Dispersão - Resíduos.....</b>	<b>4</b>

## Índice de Tabelas

<b>Tabela 1 - Tabela de Valores Simulados.....</b>	<b>5</b>
<b>Tabela 2 - Tabela de Valores Simulados Ordenada Crescentemente .....</b>	<b>5</b>

## 1. Introdução

Este relatório foi feito sobre o trabalho de grupo, da unidade curricular de métodos estatísticos.

O trabalho tem como objetivo utilizar a regressão linear simples para modelar uma variável em função de outra, utilizando os dados simulados de vendas online.

Neste relatório será analisado a correlação entre variáveis, diagramas de dispersão e os resíduos.

## 2. Dados

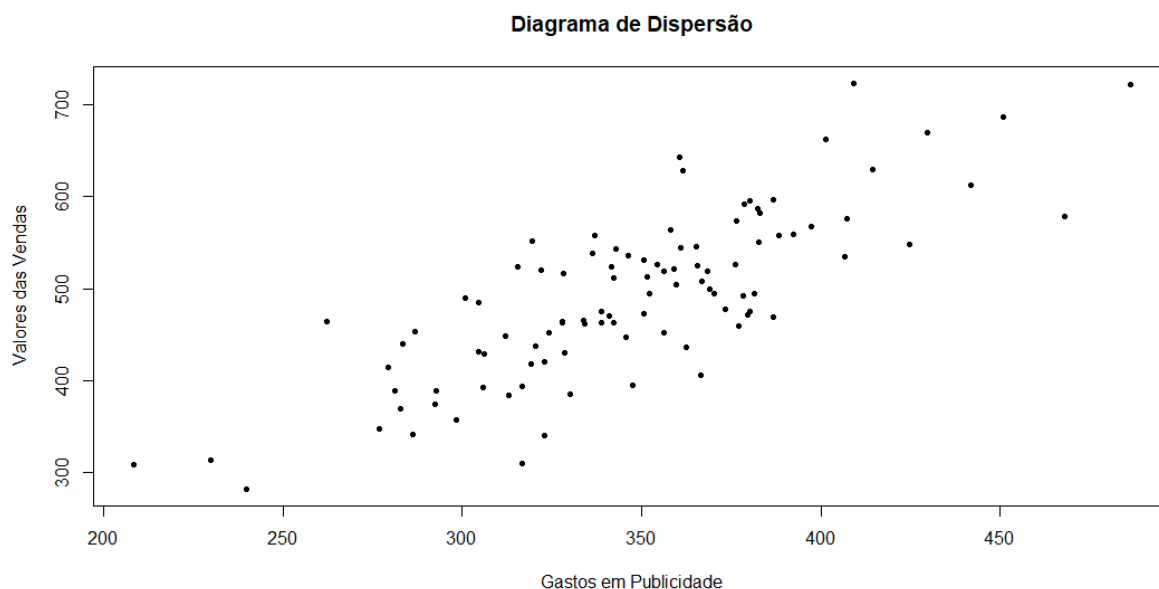
A amostra dos dados contém uma dimensão de 100 valores, as variáveis a analisar são as seguintes:

1. Publicidade (Gastos efetuados em publicidade) – Variável Quantitativas Continua
2. Vendas (Valor das vendas que ocorreram) – Variável Quantitativas Continua

Ambas as variáveis seguem uma distribuição normal onde a variável Publicidade tem uma média de 350€ e um desvio padrão de 50€, e a variável Vendas tem uma média de 500€ e um desvio padrão de 100€.

## 3. Análise das variáveis

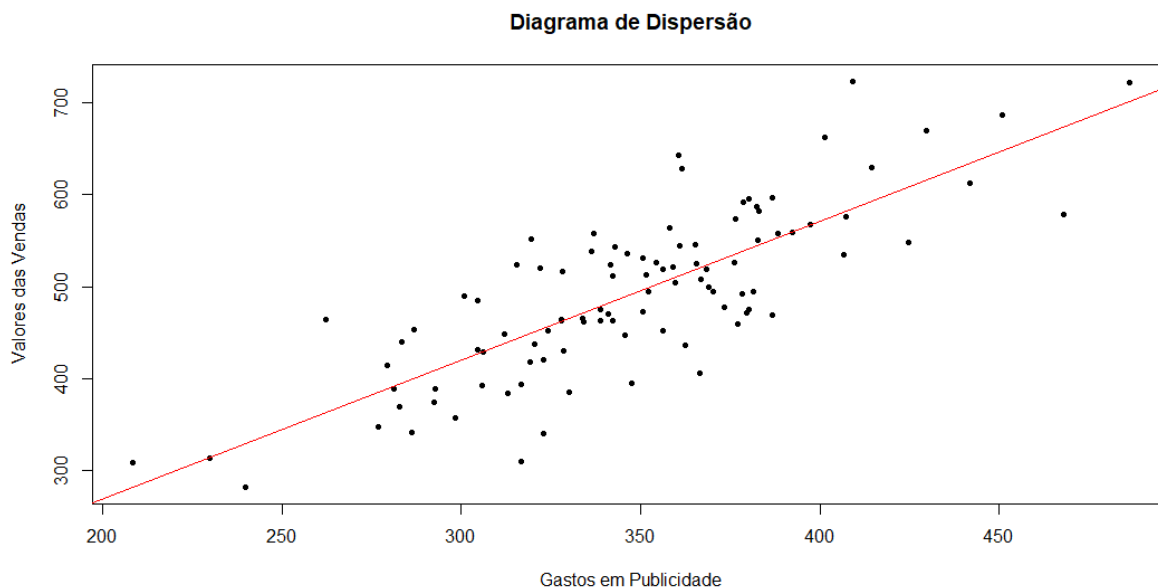
Analisado as duas variáveis obtemos o seguinte diagrama de dispersão. Podemos concluir que o coeficiente de correlação linear entre as duas variáveis é  $r_{xy} = 0.79611$ , o que mostra que existe uma correlação linear positiva moderada, mas muito perto de poder ser uma correlação linear positiva forte.



**Figura 1** - Diagrama de dispersão das duas variáveis em análise

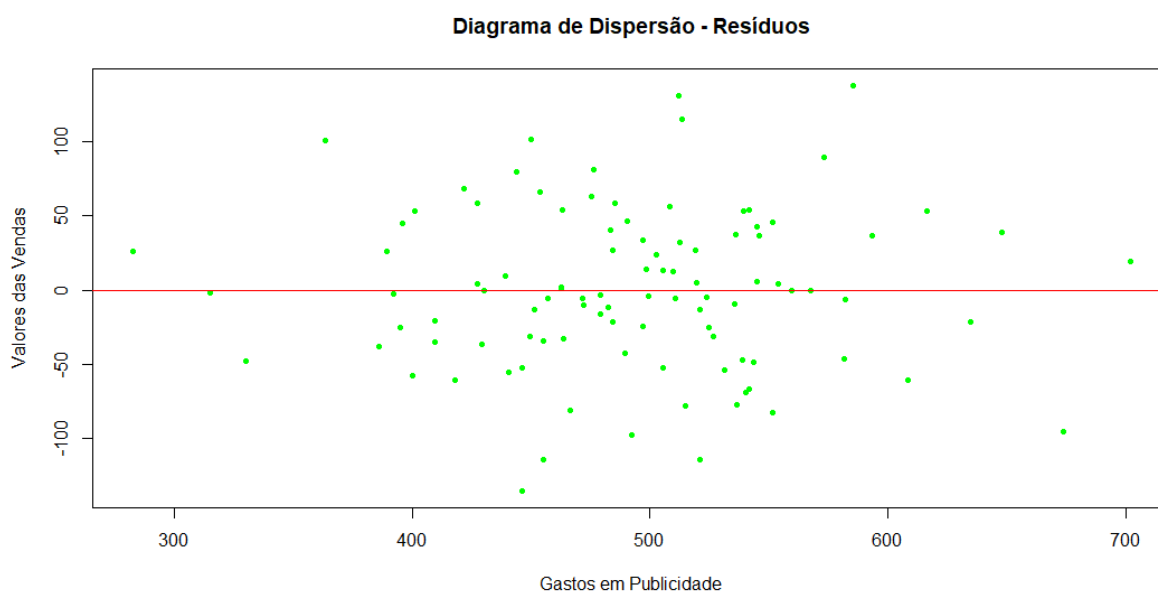
Para o modelo de Regressão Linear Simples iremos utilizar como Variável Independente a variável Publicidade = X e como Variável Dependente a variável Vendas = Y. Ao analisar a Figura 2 podemos verificar que existe uma correlação linear positiva, ou seja, a relação entre X e Y é positiva e a equação da reta é:

$$\hat{y} = -30.92490 + 1.50554x.$$



**Figura 2** - Diagrama de dispersão das duas variáveis em análise com reta regressão

Em relação aos resíduos, como podemos verificar na Figura 3, os pontos distribuem-se de forma aleatória em torno da reta que corresponde ao resíduo zero, ou seja, o modelo ajustado é bom.



**Figura 3** - Diagrama de Dispersão - Resíduos

Por fim foram simulados 5 valores possíveis de gastos em publicidade com base no modelo de regressão anterior. Os valores estão apresentados na Tabela 1.

**Tabela 1** - Tabela de Valores Simulados

n	1	2	3	4	5
Publicidade Futura	380.88083	357.28657	390.57946	387.50188	292.86617
Previsão de vendas	542.5086	506.9864	557.1104	552.4770	409.9986

**Tabela 2** - Tabela de Valores Simulados Ordenada Crescentemente

n	1	2	3	4	5
Publicidade Futura	292.86617	357.28657	380.88083	387.50188	390.57946
Previsão de vendas	409.9986	506.9864	542.5086	552.4770	557.1104

Na tabela 1 os valores encontram-se ordenados pela ordem em que foram obtidos, mas olhando para os dados de ordenados de forma crescente pelos valores gerados de publicidade futura (Tabela 2) é possível observar que, quanto maiores os gastos com a publicidade, supõem-se que maior será o valor das vendas obtido.

O que, de certa forma, faz sentido pois com mais publicidade o produto/serviço atingirá mais pessoas e, consequentemente, tornará o produto mais conhecido, o que fará com que cada vez mais pessoas queiram adquirir esse produto/serviço.

#### 4. Conclusões

Através da realização deste trabalho, podemos concluir que o coeficiente de correlação linear entre as duas variáveis Publicidade e Vendas é positivo quando comparados com os dados obtidos. Concluimos que o modelo ajustado é bom devido à distribuição dos pontos de forma aleatória e por fim, com base nos dados simulados, conseguimos fazer uma previsão de vendas futuras.

#### 5. Referências Bibliográficas

Slides disponibilizados pelos professores, na plataforma Moodle – Documento “Capítulo 1 – Estatística Descritiva”.

Slides disponibilizados pelos professores, na plataforma Moodle – Documento “Capítulo 7 – Regressão Linear Simples”.