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

João Capelas, Miguel Velez, Ricardo Sousa

[2ºL\_EI-IG-01](https://portal.ips.pt/ests/pt/hor_geral.turmas_view?pv_turma_id=602&pv_ano_lectivo=2022&pv_periodos=1&pv_periodos=2&pv_periodos=3&pv_periodos=4&pv_periodos=5&pv_periodos=6&pv_periodos=7&pv_periodos=8), 201901812

[2ºL\_EI-IG-01](https://portal.ips.pt/ests/pt/hor_geral.turmas_view?pv_turma_id=602&pv_ano_lectivo=2022&pv_periodos=1&pv_periodos=2&pv_periodos=3&pv_periodos=4&pv_periodos=5&pv_periodos=6&pv_periodos=7&pv_periodos=8), 201900678

[2ºL\_EI-IG-01](https://portal.ips.pt/ests/pt/hor_geral.turmas_view?pv_turma_id=602&pv_ano_lectivo=2022&pv_periodos=1&pv_periodos=2&pv_periodos=3&pv_periodos=4&pv_periodos=5&pv_periodos=6&pv_periodos=7&pv_periodos=8), 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

[Índice de Figuras 2](#_Toc138608967)

[Índice de Tabelas 2](#_Toc138608968)

[1. Introdução 3](#_Toc138608969)

[2. Dados 3](#_Toc138608970)

[3. Análise das variáveis 3](#_Toc138608971)

[4. Conclusões 5](#_Toc138608972)

[5. Referências bibliográficas 5](#_Toc138608973)

Índice de Figuras

[**Figura 1** - Diagrama de dispersão das duas variáveis em análise 3](#_Toc138608860)

[**Figura 2** - Diagrama de dispersão das duas variáveis em análise com reta regressão 4](#_Toc138608861)

[**Figura 3** - Diagrama de Dispersão - Resíduos 4](#_Toc138608862)

Índice de Tabelas

[**Tabela 1** -Tabela de Valores Simulados 5](#_Toc138608865)

[**Tabela 2** - Tabela de Valores Simulados Ordenada Crescentemente 5](#_Toc138608866)

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.

1. Dados

**A amostra dos dados contem 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€.

1. 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 , o que mostra que existe uma correlação linear positiva moderada, mas muito perto de poder ser uma correlação linear positiva forte.

Uma imagem com texto, captura de ecrã, diagrama, file

Descrição gerada automaticamente

**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:

Uma imagem com texto, diagrama, file, captura de ecrã

Descrição gerada automaticamente

**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.

Uma imagem com texto, captura de ecrã, diagrama, Gráfico

Descrição gerada automaticamente

**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.

1. 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. Concluímos 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.

1. 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”.