



# **Projeto: Sistema de Apoio à Decisão para a "Walmart"**

## **OBJECTIVOS:**

- Utilizar técnicas de previsão e otimização num problema do mundo real, gerindo e documentando todo o processo.
- Corresponde a 80% da nota final;

## > Grupos



- Os grupos serão constituídos de modo ideal por **4/5 alunos**.
- **Todos membros do grupo** têm de participar no desenvolvimento do projeto.
- Os grupos devem eleger um representante/líder do grupo;
- Durante as aulas, irá haver uma **avaliação do funcionamento do grupo...**

## > Autoavaliação do Projeto (A)

- No final da execução do projeto, cada grupo propõe uma Autoavaliação, nota sugerida para o projeto (de 0 a 20 valores): **A**.
- Cada projeto é avaliado pelo docente numa nota final de **P**, que pode ser igual ou diferente de **A**.
- **Exemplo de autoavaliação a inserir no relatório do projeto:** “O grupo considera que o projeto merece 16 valores, porque cumpriu este objetivos ..., mas não consegui realizar os objetivos ... e em termos de execução do projeto ao longo do tempo teve um desempenho ... .”

### Notas:

- Realizar todos os objetivos significa ter uma boa avaliação, mas não necessariamente um 18/19/20, podem existir diferentes formas, mais ou menos interessantes, de realizar os objetivos pedidos. E para além do código desenvolvido existe também a análise, o relatório, a apresentação, a execução ao longo das aulas, etc.
- Não é necessário realizar todos objetivos para ter aprovação à Unidade Curricular.

## > Autodiferenciação Individual



- Cada grupo propõe uma autodiferenciação individual, com a contribuição individual para cada membro do seu grupo.
- A autodiferenciação individual segue estas regras:
  - o **somatório individual das notas** não pode ser superior a  **$M \times P$** , onde  $M$  é o número de elementos do grupo e  $P$  é a nota final do projeto (avaliado pelo docente).
  - **não se pode reprovar um aluno** por subida de notas (para isso o grupo teria de “despedir” o aluno com antecedência);
- A autodiferenciação individual tem de ser **justificada no relatório**, com descrição do desempenho individual (ver slide sobre relatório).
- Esta autoavaliação pode ser utilizada para **diferenciar as notas, sendo a decisão final dependente do docente** (e.g., via presenças e participação nas aulas práticas, desempenho na entrega do projeto).
- Por exemplo, para uma nota  **$P$**  de 14 e 4 alunos A, B, C, D:
  - **A** trabalhou mais do que a média, **{B, C}** trabalharam um valor médio e **D** trabalhou menos: A=15, B=14, C=14, D=13.

## > Enunciado



- A **Walmart** é uma conhecida cadeia de lojas nos EUA, sendo que cada loja tem diversos departamentos (e.g., roupa, comida) que vendem produtos.
- Via o seu sistema de informação, a Walmart criou uma base de dados de vendas semanais para uma dada loja e 4 dos seus departamentos, entre 5 de fevereiro de 2010 e 26 de outubro de 2012.

## > Dados

- Os dados estão disponíveis via um ficheiro CSV: **walmart.csv**
- Na ferramenta R, os dados podem ser lidos deste modo:

```
> d=read.table("walmart.csv",sep=" ",header=TRUE)
```

```
> summary(d)
```

Date	Fuel_Price	IsHoliday	WSdep1
Length:143	Min. :2.514	Mode :logical	Min. : 30051
Class :character	1st Qu.:2.764	FALSE:133	1st Qu.: 37187
Mode :character	Median :3.290	TRUE :10	Median : 41104
	Mean :3.220		Mean : 47391
	3rd Qu.:3.594		3rd Qu.: 47252
	Max. :3.907		Max. :203670
WSdep2	WSdep3	WSdep4	
Min. :106206	Min. :45989	Min. : 93359	
1st Qu.:124488	1st Qu.:59180	1st Qu.:113069	
Median :132945	Median :62768	Median :120721	
Mean :135459	Mean :63181	Mean :120772	
3rd Qu.:145234	3rd Qu.:66590	3rd Qu.:127654	
Max. :194840	Max. :86052	Max. :148798	

## > Dados

Algumas informações adicionais sobre os dados:

- Fuel\_Price – preço de combustível na região da loja e na mesma semana
- IsHoliday – se existiu um feriado nacional na semana em causa.
- WSdep1 – valor semanal de vendas no departamento 1.
- WSdep2 – valor semanal de vendas no departamento 2.
- WSdep3 – valor semanal de vendas no departamento 3.
- WSdep4 – valor semanal de vendas no departamento 4.

## > Objetivo Previsão:



**Prever o valor semanal de vendas de cada departamento**

- Podem usar modelos univariados ou multivariados de séries temporais, bem como modelos de *machine learning*.
- No caso multivariado, podem considerar os atributos que quiserem dos dados, bem como construir outros (e.g., semana do ano, mês).
- Os modelos devem ser treinados com os dados mais antigos e testados com dados mais recentes.
- Ter em atenção o horizonte temporal, que deverá envolver uma previsão até 1 “mês” (quatro semanas) em avanço ( $H=4$  multi-step ahead predictions).



# Requisitos Previsão



- Deve ser utilizado pelo menos **um modelo** de previsão por cada **membro do grupo**. Cada membro deve ser responsável pela implementação (código), configuração (ajuste de parâmetros) e experimentação desse modelo de previsão.
- Devem usar as métricas e métodos de avaliação que considerarem mais adequados para **avaliar a qualidade** dos modelos.
- Será valorizada a qualidade do estudo comparativo de métodos, que é refletida no relatório do projeto.

## > Gestão da Loja



Para planejar um bom funcionamento da Loja, no início de cada mês (assumir que são 4 semanas), e para cada departamento da Loja, a Walmart “necessita” de definir 2 tipos de recursos:

- 1. Funcionários:** para auxiliar nas vendas (e.g. dar informações, distribuir produtos, ser “caixa”).
- 2. Produtos:** em stock e que possam ser vendidos.

## > Funcionários

Recrutados de uma bolsa de trabalhadores temporários. Cada funcionário é contratado por **um mês** e trabalha **somente num dado departamento**. Existem 3 tipos de funcionários, conforme a sua experiência:

- a)** júnior (sem experiência). Custo mensal de 6000 USD. Consegue “auxiliar”\* até 4000 produtos por semana.
- b)** normal (com experiência) Custo mensal de 8000 USD. Consegue “auxiliar” até 7000 produtos por semana.
- c)** sénior (elevada experiência). Custo mensal de 9750 USD. Consegue “auxiliar” até 9500 produtos por semana.

\* “Auxiliar” significa que somente são vendidos os produtos com procura e que podem ser auxiliados por um funcionário. Potencial total para auxiliar em  $W_{d,s}$  (e.g., um júnior no dep.1 pode ajudar a vender até 4000 produtos em qualquer das semanas do planeamento; 2 juniores no mesmo dep.1 podem ajudar em até 8000 vendas por semana no mês).

## > Produtos

- No início da semana  $s$  do mês a planear (4 semanas, logo  $s$  varia de 1 a 4), é necessário encomendar  $N_{d,s}$  produtos, que corresponde ao total de novos produtos que chegam ao departamento  $d$  na semana  $s$ .
- Cada encomenda de produtos numa semana ( $N_{d,s}$ ) tem um custo unitário (por produto) de: Dep1: 6 USD, Dep2: 8 USD, Dep3: 9 USD; Dep4: 11 USD.
- Em cada semana  $s$  e departamento  $d$ , são efetuadas todas as **vendas previstas** ( $P_{d,s}$ ) desde que haja stock e funcionários:  $P_{d,s} \leq N_{d,s}$  e  $P_{d,s} \leq W_{d,s}$ . Caso contrário, vende-se só o que for possível ser vendido: **vendas efetivas** =  $\min(N_{d,s}, W_{d,s})$ .

## > Produtos

- Seja:
  - $V_{d,s}$  o valor total de vendas efetivas no departamento **d** no final da semana **s**.
  - $S_{d,s}$  o valor em stock no departamento **d** e no início da semana **s**.
  - O próximo valor em stock (para o início da próxima semana) é dado por:  $S_{d,s+1} = S_{d,s} - V_{d,s} + N_{d,s+1}$ .
- **Cada produto** vendido tem um valor médio de venda de:
  - Dep1: 8 USD (e.g., entram em caixa 8 USD por cada unidade vendida).
  - Dep2: 10 USD.
  - Dep3: 12 USD.
  - Dep4: 16 USD.

## > **Produtos**

- Cada produto não vendido em cada semana tem custo de stock médio de:
  - Dep1: 3 USD.
  - Dep2: 5 USD.
  - Dep3: 6 USD.
  - Dep4: 8 USD.
- No final da última semana (quarta), deve-se assumir que todos produtos existentes em stock têm o custo anteriormente mencionado (e.g., 3 USD por produto em stock no Dep1).

## > **Objetivo Otimização:**

Para uma dada semana do ano (escolhida pelo gestor da loja), pretende-se definir um **plano mensal (4 semanas) de gestão da loja (funcionários e encomendas de produtos)**.

### **Objetivos a otimizar:**

**F1: Lucro** = proveitos de todas vendas – todos os custos.

**F2: “Esforço manual”** = número total de encomendas semanais por dep. e número total de funcionários contratados para o mês.

## > Exemplo de um plano:

predicted wallmart (= actual) sales:

	WSdep1	WSdep2	WSdep3	WSdep4
140	54480	159460	63584	127009
141	42221	156945	62888	124560
142	36267	146388	62768	123346
143	35283	132156	60279	117375

hired workers (type x dept):

	[,1]	[,2]	[,3]	[,4]
[1,]	5	4	3	2
[2,]	6	5	4	3
[3,]	7	6	5	4

total cost workers: 442500

effort > total number of workers: 54

worker max. support per dept: 128500 108000 87500 67000

product orders (week x dept):

	[,1]	[,2]	[,3]	[,4]
[1,]	61662	78292	56434	24182
[2,]	0	0	0	0
[3,]	12985	55403	69133	37167
[4,]	39924	75160	62131	99708

total cost of orders: 5819175

effort > total number of orders per dep and week: 12

Elementos do  
plano mensal

(solução concreta que o  
método de otimização  
poderá gerar)



sales (week x dept):

	[,1]	[,2]	[,3]	[,4]
--	------	------	------	------

[1,]	54480	78292	56434	24182
------	-------	-------	-------	-------

[2,]	7182	0	0	0
------	------	---	---	---

[3,]	12985	55403	62768	37167
------	-------	-------	-------	-------

[4,]	35283	75160	60279	67000
------	-------	-------	-------	-------

Não existe aqui produto para ser vendido

sales in USD (week x dept):

	[,1]	[,2]	[,3]	[,4]
--	------	------	------	------

[1,]	435840	782920	677208	386912
------	--------	--------	--------	--------

[2,]	57456	0	0	0
------	-------	---	---	---

[3,]	103880	554030	753216	594672
------	--------	--------	--------	--------

[4,]	282264	751600	723348	1072000
------	--------	--------	--------	---------

total revenue sales: 7175346

stock (week x dept):

	[,1]	[,2]	[,3]	[,4]
--	------	------	------	------

[1,]	7182	0	0	0
------	------	---	---	---

[2,]	0	0	0	0
------	---	---	---	---

[3,]	0	0	6365	0
------	---	---	------	---

[4,]	4641	0	8217	32708
------	------	---	------	-------

cost of stock in USD (week x dept:)

	[,1]	[,2]	[,3]	[,4]
--	------	------	------	------

[1,]	21546	0	0	0
------	-------	---	---	---

[2,]	0	0	0	0
------	---	---	---	---

[3,]	0	0	38190	0
------	---	---	-------	---

[4,]	13923	0	49302	261664
------	-------	---	-------	--------

total stock cost: 384625

total costs: 442500 + 5819175 + 384625 = 6646300

>> F1 (month profit)= 529046

>> F2 (month effort)= 66

# Requisitos otimização:

- Para a otimização, assumir: valores previstos de vendas de produtos para as próximas 4 semanas (multi-step ahead, **melhor**) OU usar valores reais;
- Assumir 2 possibilidades de otimização:
  - **O1**: Maximizar o lucro da empresa para um mês (**F1**).
  - **O2**: Multiobjetivo: maximizar o lucro da empresa (**F1**) e minimizar o esforço (**F2**).
- Deve ser utilizado pelo menos um modelo de otimização por cada membro do grupo. Cada membro deve ser responsável pela implementação (código), configuração (ajuste de parâmetros) e experimentação desse modelo de otimização.

# Valorização



- **Qualidade do estudo** sobre a **previsão e otimização** (medida no relatório).
- Maior valorização usarem previsões e não valores reais na otimização, se testarem e compararem diversos modelos de previsão e otimização (métodos explicados nas aulas, outros métodos de outros pacotes do R ou de outras ferramentas), estudar o efeito de alterar algumas das variáveis (e.g., data em que se faz a simulação, valores de custo).
- Na otimização, pode ser acrescentado mais realismo ao problema, criando outras versões a otimizar.
- Sistema desenvolvido, incluindo **interface** em modo de consola ou **gráfica** (e.g., via [shiny.rstudio.com](https://shiny.rstudio.com)) para demonstrar o sistema completo de Análise Inteligente de Dados (previsão e otimização) a funcionar.

## > Notas:

- Não existe uma única forma ideal de resolver o projeto, é esperado que cada grupo execute um projeto diferente;
- Aceito alguma flexibilidade/originalidade em cada grupo, desde que **justifiquem as decisões tomadas**;
- Não necessitam de executar tudo o que é pedido para ter nota positiva.

## > Execução

- A partir de hoje até data de entrega, apoio durante aulas.
- Durante as aulas será efetuada uma avaliação preliminar (não formal, com um peso reduzido) a cada grupo e projeto (convém estarem presentes todos elementos do grupo).
- A partir do dia 21/3/2024, todos grupos devem mostrar ao docente um “semanário” de projeto (via Googledocs, overleaf, etc.), que deve incluir para cada semana uma entrada com:
  - Resumo do que foi **realizado nessa semana para o projeto**. Indicação de que **reuniões de trabalho** tiveram lugar (com o grupo todo) e sua duração.
  - Para **cada elemento do projeto**: o que fez em termos individuais e quantas horas dedicou ao projeto.

# Exemplo de um semanário:

Trabalho de grupo	
<b>Resumo do trabalho realizado:</b> exploração e aplicação das técnicas Rolling Window e Growing Window aos modelos Arima, Holt-Winters, NN, ETS (package forecast) e MLPE e Random Forest (package rminer)	
Número de reuniões realizadas: 2	
<ul style="list-style-type: none"> <li>Descrição da reunião nº1</li> </ul>	Dia: 05/04/2023
	Duração: 2horas
	Objetivo(s): <ul style="list-style-type: none"> <li>Análise da script fornecida pelo docente associada às técnicas de Growing e Rolling Window.</li> <li>Divisão pelos elementos do grupo dos diferentes modelos aplicados ao Rolling Window - biblioteca Forecast e biblioteca rminer.</li> <li>Divisão pelos elementos do grupo dos diferentes modelos aplicados ao Growing Window - biblioteca Forecast e biblioteca rminer.</li> </ul>
<ul style="list-style-type: none"> <li>Descrição da reunião nº2</li> </ul>	Dia: 11/04/2023
	Duração: 1 hora
	Objetivo(s): <ul style="list-style-type: none"> <li>Análise do modelos estudados pelo grupo;</li> <li>Discussão dos resultados obtidos nos modelos estudados;</li> <li>Debate das dúvidas dos elementos do grupo;</li> </ul>

Trabalho individual		
Elemento	Trabalho / Tarefas realizadas	Horas dedicadas ao projeto
Vasco Pereira A92927	→ Exploração da documentação da biblioteca Rminer, para implementação do modelo Random Forest. → Análise univariada das vendas de Stella e Bud, com utilização de Growing Window e Rolling Window, dos modelos da biblioteca rminer. → Compilação das receitas dos vários elementos. → Análise de resultados.	3h
Daniela Lomba A92969	→ Análise univariada para a cerveja Stella e Bud, com utilização da técnica Rolling Window, do modelo convencional Arima e Ets. → Análise e interpretação dos resultados obtidos.	2h
Nuno Amorim A90156	→ Análise univariada para a cerveja Stella e Bud, com utilização da técnica Growing Window, dos modelos da biblioteca forecast. → Análise e interpretação dos resultados obtidos.	2h
Raquel Silva A92966	→ Realização da análise univariada utilizando a técnica Rolling Window para cerveja Stella e Bud do modelo convencional Holt-Winters e NN (forecast). → Avaliação e análise dos resultados obtidos.	2h

## Dúvidas:

- ☒ Step-jumps, não-deveríamos por 7? Porque avançamos uma semana
- ☒ Na script de growing-window do rminer o modelo não chega a fazer previsões para a última semana.

## > Entrega

- Dia **23/5/2024** (quinta, dentro do horário das aulas e conforme horário a disponibilizar no e-learning).
- Cada grupo (todos elementos presentes) deve efetuar uma **apresentação** do trabalho realizado, respondendo a questões.
- Entrega de 1) **relatório** (em pdf), 2) **todo código R desenvolvido**, 3) **slides da apresentação (em formato pdf, opcional)** => um único ficheiro **.zip** a depositar no sistema elearning.

# Video de Demonstração

- Deve ser **gravado um vídeo** (até máximo de 5 minutos, idealmente com narração) onde se demonstra o sistema desenvolvido (em R e/ou outro), sendo efetuado o upload no **youtube**.
- Incluir no relatório o link (ou links) do vídeo **youtube**.
- Existe diverso software gratuito de captura de ecrã para vídeo, por exemplo, fazer a pesquisa google:  
**screen capture video**



# > Relatório

- O relatório deve explicar aquilo que foi realizado:
  1. **Introdução** (breve descrição do que foi feito)
  2. **Execução do Projeto:** Os alunos que constituem o grupo, planeamento (que tipo de tarefas cada um executou ao longo do tempo), como funcionou o grupo, qual a autoavaliação (nota proposta para o projeto e diferenciação de notas). Ter 1 página por elemento do grupo, onde se descreve o que esse elemento fez em termos do projeto (e.g., esforço, tarefas, código, testes).
  3. **Objetivo de Previsão** (o que foi feito e resultados)
  4. **Objetivo de Otimização** (o que foi feito e resultados)
  5. **Demonstração do Sistema Desenvolvido** (1 frase com o link [youtube](#))
  6. **Conclusões** (breve apreciação do que foi realizado).
    - **Bibliografia** (opcional)
    - **Anexos** (opcional)

Tentar ter um máximo a rondar as 20 a 40 páginas de corpo de documento (sem capa, índices, bibliografia e anexos). Escrita direta, 25 sem “palha”!