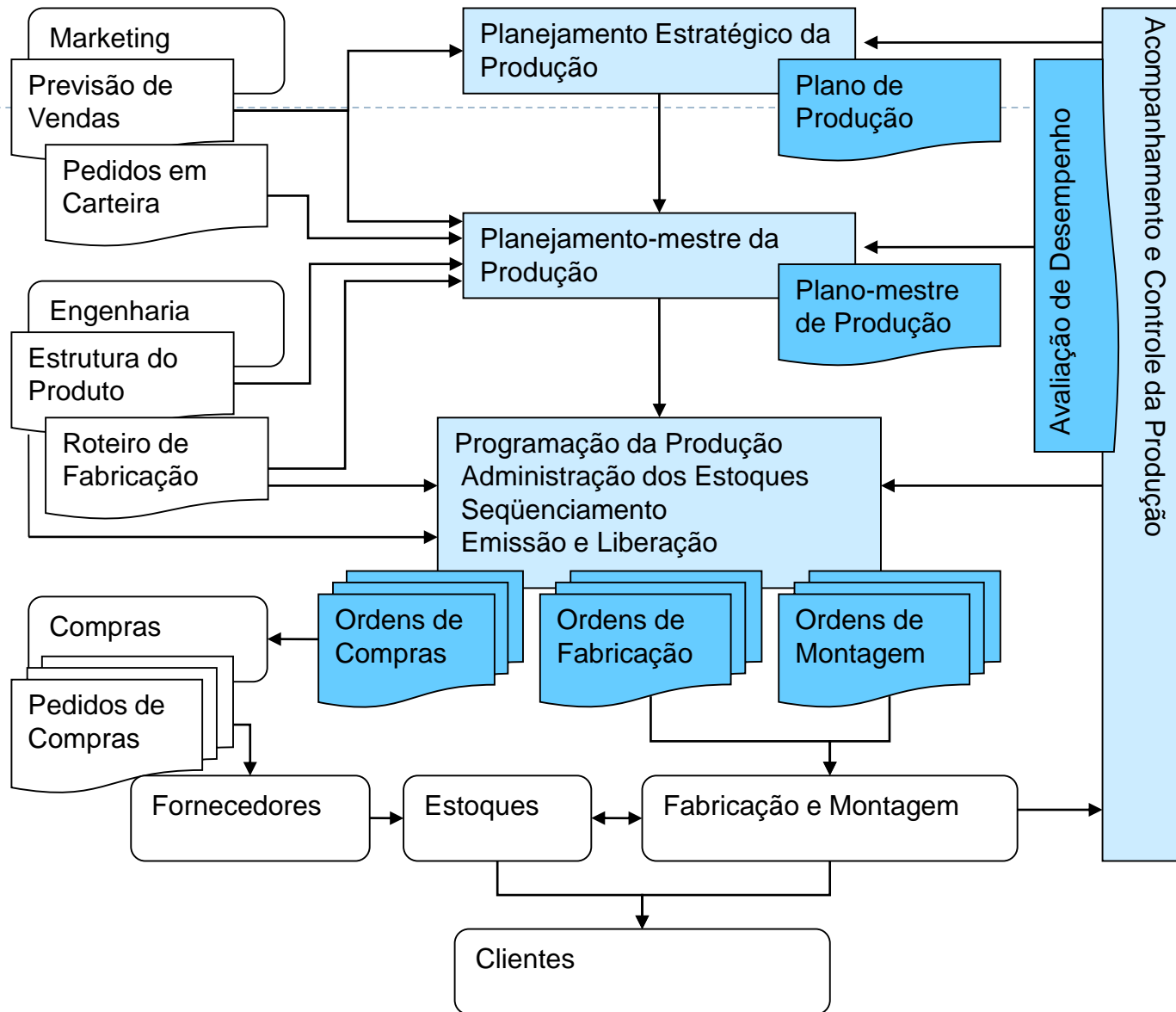


# GESTÃO DA PRODUÇÃO

## *Introdução à Previsão de Demanda*

**UFES**

Prof. Dr. Joao Ferreira Netto



### *Análise de regressão*

*Ferramenta de inferência estatística que permite construir um modelo que identifica as relações entre variáveis independentes  $X_i$  e uma variável de resposta  $Y$ .*

### *Regressão linear múltipla*

*Função (modelo) linear que relaciona um conjunto de variáveis independente  $X_i$  com uma única variável de resposta  $Y$ .*

**Exercício:** Foram obtidos os custos operacionais mensais para um período de 19 meses de uma fábrica que produz três produtos A, B e C.

Deseja-se obter um modelo de regressão que melhor represente o custo operacional mensal da fábrica em função do número de unidades produzidas A, B e C.

*Modelo:* 
$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \varepsilon \quad \varepsilon \sim N(0, \sigma)$$

*Estatísticas:* 
$$\hat{y}_i = a + b_1 x_1 + b_2 x_2 + \dots$$

*O Excel calcula os valores da constante e de  $B_1, B_2, \dots, B_n$  para que sejam feitas as previsões a partir de tal equação da forma mais precisa possível (buscando minimizar a soma dos erros quadráticos).*

*Abra a aba “Fábrica” no arquivo “regressão\_múltipla\_alunos.xls” disponibilizado.*

Os valores dos custos operacionais mensais (variáveis dependentes) e das quantidades produzidas dos produtos A, B e C (variáveis independentes) estão na tabela.

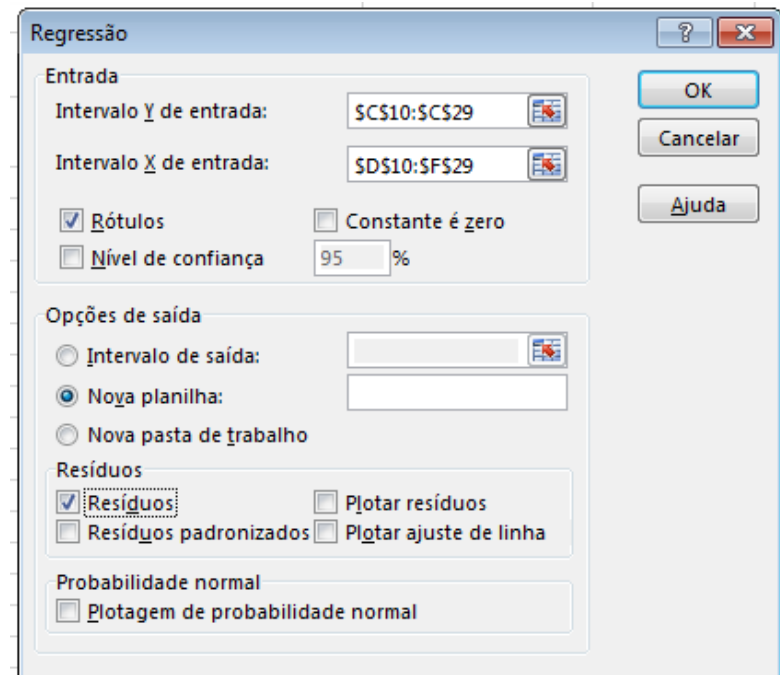
Deseja-se encontrar a melhor previsão para o custo operacional mensal no formato:

$$\begin{aligned} \text{Custo operacional mensal} = & \text{constante} + \\ & B1 * (\text{Unidades produzidas de A}) + \\ & B2 * (\text{Unidades produzidas de B}) + \\ & B3 * (\text{Unidades produzidas de C}) \end{aligned}$$

Mês	Custo operacional mensal	Unidades produzidas		
		A	B	C
1	R\$ 44.439,00	515	541	928
2	R\$ 43.936,00	929	692	711
3	R\$ 44.464,00	800	710	824
4	R\$ 41.533,00	979	675	758
5	R\$ 46.343,00	1165	1147	635
6	R\$ 44.922,00	651	939	901
7	R\$ 43.203,00	847	755	580
8	R\$ 43.000,00	942	908	589
9	R\$ 40.967,00	630	738	682
10	R\$ 48.582,00	1113	1175	1050
11	R\$ 45.003,00	1086	1075	984
12	R\$ 44.303,00	843	640	828
13	R\$ 42.070,00	500	752	708
14	R\$ 44.353,00	813	989	804
15	R\$ 45.968,00	1190	823	904
16	R\$ 47.781,00	1200	1108	1120
17	R\$ 43.202,00	731	590	1065
18	R\$ 44.074,00	1089	607	1132
19	R\$ 44.610,00	786	513	839

*Para fazer a regressão, vá até a aba de ferramentas do Excel “Dados” e selecione a opção “Análise de Dados” (caso não esteja disponível, instale o suplemento).*

- *O intervalo de entrada Y, C10:C29, contém as variáveis dependentes ou dados (incluindo o rótulo custo) que se deseja prever.*
- *O intervalo de entrada X, D10: F2922, contém os dados ou variáveis independentes (incluindo os rótulos) que se deseja usar na previsão (o Excel tem um limite de 15 variáveis independentes)*
- *Uma vez que tanto o input X quanto Y incluem os rótulos, clique na caixa Rótulos.*
- *Escolhe-se uma planilha para a saída dos resultados.*
- *Marcar a opção Resíduos faz com que Excel liste para cada observação a previsão através da fórmula gerada e os erros de previsão (custo observado menos o custo previsto).*



Encontramos na coluna Coeficientes que a melhor equação a ser utilizada para prever o custo mensal é:

$$\text{Custo mensal previsto} = 35.102,90 + 2,07 * A + 4,18 * B + 4,79 * C$$

	Coeficientes
Interseção	35102,90045
A	2,065953296
B	4,176355531
C	4,790641037

A pergunta natural é: nossas variáveis independentes são úteis para prever o custo mensal?

R: Qualquer variável independente com um p-value (ver coluna E) menor ou igual a 0,15 é considerada útil para a previsão da variável dependente.

valor-P
0,00000000000611
0,23372682031886
0,02528778545180
0,01722264282113



- *Os p-values indicam que A não acrescenta poder preditivo a B e C, o que significa que, se conhecemos B e C, podemos prever o custo operacional mensal tão bem quanto poderíamos se incluíssemos A como uma variável independente.*
- *Portanto, podemos optar por excluir A como uma variável independente e usar apenas B e C.*
- *Copie os dados sem as quantidades de A em outra planilha e refaça o processo de regressão múltipla.*

*Observa-se que B e C possuem p-values muito baixos (0,002 e 0,007, respectivamente). Estes valores dizem que ambas as variáveis independentes têm poder preditivo útil.*

*Usando os novos coeficientes da coluna B, agora pode-se prever o custo operacional mensal usando a equação:*

$$\text{Custo operacional mensal previsto} = 35.475,3 + 5,32 * B + 5,42 * C$$

	<i>Coeficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>
Interseção	35475,30255	1842,860853	19,25012543	1,72346E-12
B	5,320968077	1,429095476	3,72331182	0,001849065
C	5,417137848	1,745311646	3,103822668	0,006825007

### ***Quão precisas são as previsões do custo mensal de unidades produzidas?***

- *Na célula B5 da planilha, observa-se que  $R^2$  é igual a 0,61. Tal valor de  $R^2$  mostra que, em conjunto, B e C explicam 61% da variação dos custos operacionais mensais.*

*Obs: na regressão original, que incluía A como uma variável independente,  $R^2$  é igual a 0,65, indicando que a adição de A como uma variável independente explica apenas 4% da variação nos custos operacionais mensais. Ter uma diferença tão pequena é consistente com a exclusão de A como variável independente.*

- *Na célula B7 da planilha, observa-se que o erro padrão para a regressão com B e C como variáveis independentes é 1.274.*
- *Espera-se que cerca de 68% das previsões de regressão múltipla sejam exatas dentro de um erro padrão, e 95% das previsões de regressão múltipla sejam exatas dentro de dois desvios-padrão.*
- *Qualquer previsão que difere do valor real em mais de dois desvios-padrão é considerada um caso isolado. Assim, se o nosso custo operacional previsto é em erro por mais de \$ 2.548 ( $2 * 1274$ ), considera-se que a observação é um outlier.*

- Na tabela “Resultados de Resíduos” apresenta-se para cada observação do custo, o resíduo previsto, que equivale ao custo real menos o custo previsto.
- Pela primeira observação, por exemplo, pode-se prever um custo de US \$ 43.381,10 e o residual de \$ 1.057,95 indica que a previsão errou o custo real por \$ 1.057,95.

RESULTADOS DE RESÍDUOS		
Observação	Previsto(a) Custo operacional mensal	Resíduos
1	43381,05021	1057,949794
2	43008,99747	927,002527
3	43716,91148	747,0885248
4	43173,14649	-1640,146495
5	45018,33547	1324,664528
6	45352,53278	-430,5327792
7	42634,5734	568,4265963
8	43497,43576	-497,4357601
9	43096,66501	-2129,665007
10	47415,43478	1166,565215
11	46525,80688	-1522,806879
12	43366,11226	936,8877388
13	43312,00414	-1242,004144
14	45093,11881	-740,1188117
15	44751,5519	1216,448104
16	47438,12957	342,8704271
17	44383,92553	-1181,925527
18	44837,33022	-763,3302206
19	42749,93783	1860,062168

### ***Incorporação de fatores qualitativos na regressão múltipla***

- *Em muitas situações, as variáveis independentes não podem ser facilmente quantificadas.*
- *Por esse motivo, vamos olhar para formas de incorporar fatores qualitativos, tais como a sazonalidade, gênero ou o partido de um candidato presidencial, em uma análise de regressão múltipla.*

### ***Como posso prever as vendas trimestrais de automóveis dos EUA?***

*Suponha que se deseja prever as vendas trimestrais de automóveis dos EUA para se entender como o trimestre do ano influencia nas vendas de automóveis. São utilizados os dados da aba “Vendas de carros” (as vendas são listadas em milhares de carros).*

### **Incorporação de fatores qualitativos na regressão múltipla**

- *Você pode ser tentado a definir uma variável independente que seja igual a 1 no primeiro trimestre, a 2 durante o segundo trimestre, e assim por diante.*
- *Infelizmente, essa abordagem forçaria o quarto trimestre a ter quatro vezes o efeito do primeiro trimestre, o que pode não ser verdade.*
- **O trimestre do ano é uma variável independente qualitativa.**
- *Para modelar uma variável independente qualitativa, deve-se criar uma variável independente (chamada aqui de variável dummy) para todos os possíveis valores da variável qualitativa;*
- *Uma das variáveis pode ser deixada de fora. Neste exemplo, optou-se por deixar de fora o quarto trimestre.*
- *As variáveis dummy mostram que valor da variável qualitativa ocorre. Assim, obtém-se uma variável dummy para o Trimestre 1, 2 e 3 com as seguintes propriedades:*
  - **Variável dummy do trimestre 1 igual a 1 se o trimestre é o trimestre 1 e 0 caso contrário;**
  - **Variável dummy do trimestre 2 igual a 1 se o trimestre é o trimestre 2 e 0 caso contrário;**
  - **Variável dummy do trimestre 3 igual a 1 se o trimestre é o trimestre 3 e 0 caso contrário.**

## Incorporação de fatores qualitativos na regressão múltipla

	A	B	C	D	E	F	G	H	I
1	Dados Históricos								
2									
3	Ano	Trimestre	Vendas	PIB	Desemprego	Inflação	T1	T2	T3
4	79	1	Vendas	2541	5,9	9,4	T1	T2	T3
5	79	2	2910	2640	5,7	9,4	0	1	0
6	79	3	2562	2595	5,9	9,7	0	0	1
7	79	4	2385	2701	6	11,9	0	0	0
8	80	1	2520	2785	6,2	13,4	1	0	0
9	80	2	2142	2509	7,3	9,6	0	1	0
10	80	3	2130	2570	7,7	9,2	0	0	1
11	80	4	2190	2667	7,4	13,6	0	0	0
12	81	1	2370	2878	7,4	14,4	1	0	0
13	81	2	2208	2835	7,4	15,3	0	1	0
14	81	3	2196	2897	7,4	15,1	0	0	1
15	81	4	1758	2744	8,3	11,8	0	0	0
16	82	1	1944	2582	8,8	12,8	1	0	0
17	82	2	2094	2613	9,4	12,4	0	1	0
18	82	3	1911	2529	10	9,3	0	0	1
19	82	4	2031	2544	10,7	7,9	0	0	0
20	83	1	2046	2633	10,4	7,8	1	0	0
21	83	2	2502	2878	10,1	8,4	0	1	0
22	83	3	2238	3051	9,4	9,1	0	0	1
23	83	4	2391	3274	8,5	8,8	0	0	0
24	84	1	2586	3584	7,9	9,2	1	0	0
25	84	2	2896	3774	7,5	9,6	0	1	0
26	84	3	2448	3861	7,5	10,3	0	0	1

## Incorporação de fatores qualitativos na regressão múltipla

- Além de sazonalidade, podem ser consideradas variáveis macroeconômicas, como o produto interno bruto (PIB, em bilhões de dólares), taxas de juros e as taxas de desemprego para prever vendas de automóveis.
- Suponha, por exemplo, que estamos tentando prever as vendas para o segundo trimestre de 1979. Como valores de PIB, taxa de juros e taxa de desemprego não são conhecidos no início do segundo trimestre 1979, os dados não podem ser utilizados para prever as vendas de automóveis do segundo trimestre de 1979.
- Ao invés disso, vamos utilizar os valores do PIB, taxas de juros e taxas de desemprego defasadas de um trimestre para prever as vendas de automóveis.

Dados Históricos											
Ano	Trimestre	Vendas	PIB	Desemprego	Inflação	T1	T2	T3	PIB Defas	Desemprego Defas	Inflação Defas.
79	1	Vendas	2541	5,9	9,4	T1	T2	T3	PIB Defas	Desemprego Defas	Inflação Defas.
79	2	2910	2640	5,7	9,4	0	1	0	2541	5,9	9,4
79	3	2562	2595	5,9	9,7	0	0	1	2640	5,7	9,4
79	4	2385	2701	6	11,9	0	0	0	2595	5,9	9,7



## Incorporação de fatores qualitativos na regressão múltipla

- Agora podemos executar a regressão múltipla pela Análise de Dados/ Regressão;
- Utiliza-se C4: C26 como o intervalo de entrada Y, G4: L26 como o intervalo de entrada X;
- Selecione a opção Rótulos e também selecione a opção Resíduos.
- Qual a equação obtida?
- Vamos interpretar os coeficientes!

	Coeficientes	Erro padrão	Stat t	valor-P
Interseção	3156,83203	493,4577588	6,397370339	1,1997E-05
T1	164,0664875	115,8189309	1,416577464	0,17703895
T2	303,9970934	111,9318641	2,715912005	0,015941658
T3	65,33086962	112,0265637	0,583173021	0,568441843
PIB Defas	0,252374234	0,119241827	2,11649084	0,051432389
Desemprego	-104,5168927	27,75125677	-3,766203945	0,00186798
Inflação De	-81,09860654	17,93418199	-4,522013136	0,000405138

- Tabela ANOVA: A única quantidade de interesse na tabela de análise de variância é a significância (0,00128). Esta medida implica que só há 1,28 chances em mil de que as variáveis independentes sejam inúteis na previsão de vendas de automóveis.

*Quando perguntado sobre que fatores são determinantes nas eleições presidenciais, o especialista em eleições James Carville disse: "It's the economy, stupid". O economista de Yale Roy Fair mostrou que Carville está correto em achar que a economia é um fator importante nos resultados eleitorais.*

- ***Deseja-se obter um modelo de regressão que melhor represente a votação obtida pelo partido do governo nas eleições presidenciais americanas.***
- *Busca-se prever a porcentagem do partido no poder do voto bipartidário usando as seguintes variáveis independentes:*
- *Partido no poder.*
- *A porcentagem de crescimento do PIB durante os primeiros nove meses do ano da eleição.*
- *Valor absoluto da taxa de inflação durante os primeiros nove meses do ano da eleição (valores absolutos)*
- *Número de trimestres durante os últimos quatro anos em que o crescimento econômico foi significativo (> 3,2%)*
- *Tempo em que o partido está no cargo. Variável é igual a 0 para denotar um mandato, 1 para dois mandatos, 1,25 para três mandatos, 1,5 por quatro mandatos, e 1,75 por pelo menos cinco. Esta definição implica que após o primeiro mandato cada outro tem menos influência sobre o resultado das eleições.*
- *A eleição ocorre durante a guerra?*
- *O atual presidente disputa a reeleição?*