

# B2W LABs

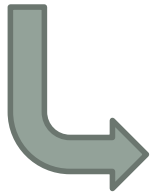
---

## Pricing Challenge

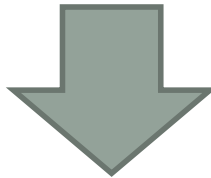
William T. Maruyama

# Resumo do problema

Relação de preço e demanda no mercado



Ajuste/Otimização



Criação de modelo preditivo da quantidade de produtos vendidos a partir do preço.

# Etapas desenvolvidas

## **1** – Análise dos dados de vendas

- Sumarização
- Agrupamento
- Visão geral receita
- Visão geral quantidade vendida

## **2** – Modelo preditivo: Dado um preço, qual a quantidade vendida?

## **3** – Análise concorrência X vendas

- Sumarização
- Análise pelo gráfico
- Agrupamento

# Dados

## Sales

- 9 produtos (P1, P2, ..., P9)
- Data da ordem de venda
- Quantidade vendida
- Receita

Categórico

Numérico

## Comp\_Prices

- 6 concorrentes (C1, C2, ..., C6)
- 9 produtos (P1, P2, ..., P9)
- Data e h da extração dos dados (2x ao dia)
- Preço do produto
- Tipo de pagamento (1: a prazo; 2: à vista)

Categórico

Numérico

Sumarização: Métricas calculadas em cada coluna do conjunto para cada Produto

PROD_ID	DATE_ORDER	QTY_ORDER	REVENUE
P1 :4091	2015-03-18: 65	Min. : 1.00	Min. : 1146
	2015-02-26: 62	1st Qu.: 1.00	1st Qu.: 1399
	2015-03-04: 60	Median : 1.00	Median : 1499
	2015-09-02: 57	Mean : 1.02	Mean : 1470
	2015-03-12: 54	3rd Qu.: 1.00	3rd Qu.: 1499
	2015-03-11: 53	Max. : 20.00	Max. : 25888
	(Other) :3740		
PROD_ID	DATE_ORDER	QTY_ORDER	REVENUE
P2 :63358	2015-08-19: 3549	Min. : 1.000	Min. : 494.1
	2015-05-27: 3414	1st Qu.: 1.000	1st Qu.: 629.1
	2015-08-20: 1938	Median : 1.000	Median : 679.0
	2015-05-26: 1774	Mean : 1.071	Mean : 726.3
	2015-03-18: 1492	3rd Qu.: 1.000	3rd Qu.: 729.0
	2015-08-18: 1360	Max. : 50.000	Max. : 27907.0
	(Other) :49831		
PROD_ID	DATE_ORDER	QTY_ORDER	REVENUE
P3 :2899	2015-03-11: 93	Min. :1.000	Min. :1009
	2015-07-06: 54	1st Qu.:1.000	1st Qu.:1099
	2015-07-20: 52	Median :1.000	Median :1288
	2015-03-10: 51	Mean :1.017	Mean :1281
	2015-07-21: 44	3rd Qu.:1.000	3rd Qu.:1358
	2015-03-12: 43	Max. :6.000	Max. :6588
	(Other) :2562		

# 1 – Análise dos dados de vendas

	PROD_ID	DATE_ORDER	QTY_ORDER	REVENUE	
<b>P4</b>	:14656	2015-09-23:	1170	Min. : 1.000	Min. : 265.8
		2015-09-24:	954	1st Qu.: 1.000	1st Qu.: 474.1
		2015-05-04:	879	Median : 1.000	Median : 499.0
		2015-05-15:	443	Mean : 1.181	Mean : 577.0
		2015-09-22:	426	3rd Qu.: 1.000	3rd Qu.: 512.1
		2015-05-16:	339	Max. :109.000	Max. :53192.0
		(Other)	:10445		

	PROD_ID	DATE_ORDER	QTY_ORDER	REVENUE	
<b>P5</b>	:19775	2015-08-19:	1288	Min. : 1.000	Min. : 561.8
		2015-09-02:	1055	1st Qu.: 1.000	1st Qu.: 749.0
		2015-09-01:	902	Median : 1.000	Median : 749.0
		2015-08-20:	766	Mean : 1.065	Mean : 804.4
		2015-09-03:	680	3rd Qu.: 1.000	3rd Qu.: 779.0
		2015-10-01:	564	Max. :41.000	Max. :27452.4
		(Other)	:14520		

	PROD_ID	DATE_ORDER	QTY_ORDER	REVENUE	
<b>P6</b>	:4102	2015-09-23:	371	Min. : 1.000	Min. : 1133
		2015-02-04:	362	1st Qu.: 1.000	1st Qu.: 1500
		2015-02-03:	280	Median : 1.000	Median : 1698
		2015-09-24:	206	Mean : 1.033	Mean : 1726
		2015-02-27:	155	3rd Qu.: 1.000	3rd Qu.: 1787
		2015-03-04:	135	Max. :24.000	Max. :33576

# 1 – Análise dos dados de vendas

	PROD_ID	DATE_ORDER	QTY_ORDER	REVENUE
<b>P7</b>	:195939	2015-08-19: 4930	Min. : 1.000	Min. : 374.5
		2015-08-20: 3276	1st Qu.: 1.000	1st Qu.: 739.0
		2015-05-27: 3239	Median : 1.000	Median : 749.0
		2015-09-02: 3065	Mean : 1.081	Mean : 822.9
		2015-09-01: 2728	3rd Qu.: 1.000	3rd Qu.: 799.0
		2015-03-03: 2485	Max. : 500.000	Max. : 368750.0
		(Other) :176216		

	PROD_ID	DATE_ORDER	QTY_ORDER	REVENUE
<b>P8</b>	:27840	2015-08-19: 1616	Min. : 1.000	Min. : 229.0
		2015-08-21: 1400	1st Qu.: 1.000	1st Qu.: 379.0
		2015-08-20: 1031	Median : 1.000	Median : 394.9
		2015-09-01: 947	Mean : 1.071	Mean : 436.7
		2015-09-02: 907	3rd Qu.: 1.000	3rd Qu.: 449.0
		2015-10-14: 839	Max. : 40.000	Max. : 15160.0
		(Other) :21100		

	PROD_ID	DATE_ORDER	QTY_ORDER	REVENUE
<b>P9</b>	:18431	2015-08-19: 858	Min. : 1.000	Min. : 307.0
		2015-08-21: 807	1st Qu.: 1.000	1st Qu.: 379.0
		2015-09-02: 789	Median : 1.000	Median : 394.0
		2015-09-01: 709	Mean : 1.067	Mean : 432.9
		2015-08-20: 684	3rd Qu.: 1.000	3rd Qu.: 434.7
		2015-08-18: 648	Max. : 40.000	Max. : 13795.6
		(Other) :13936		

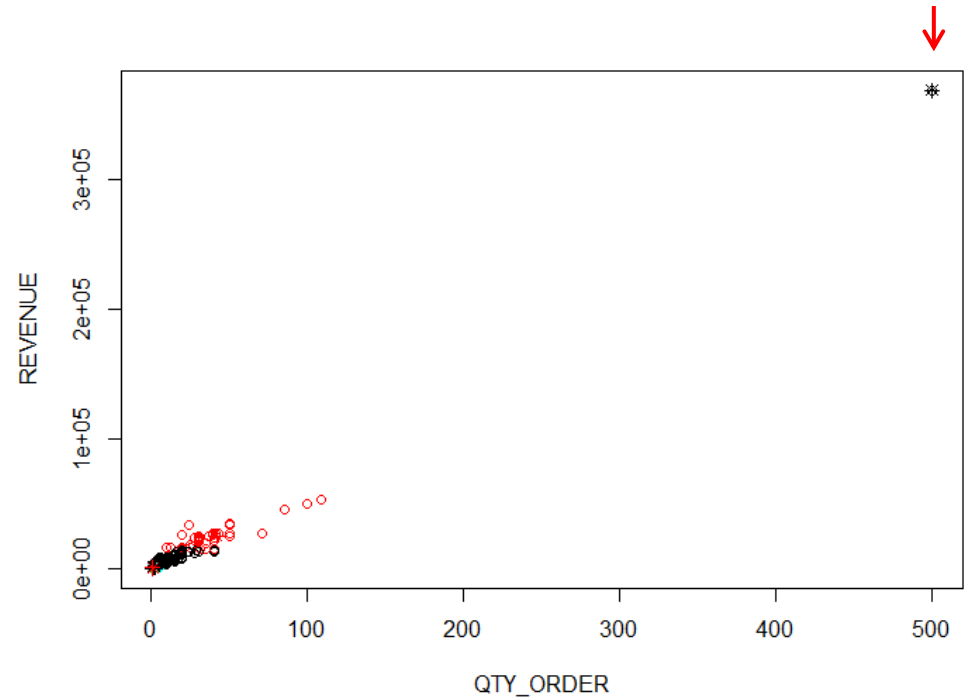
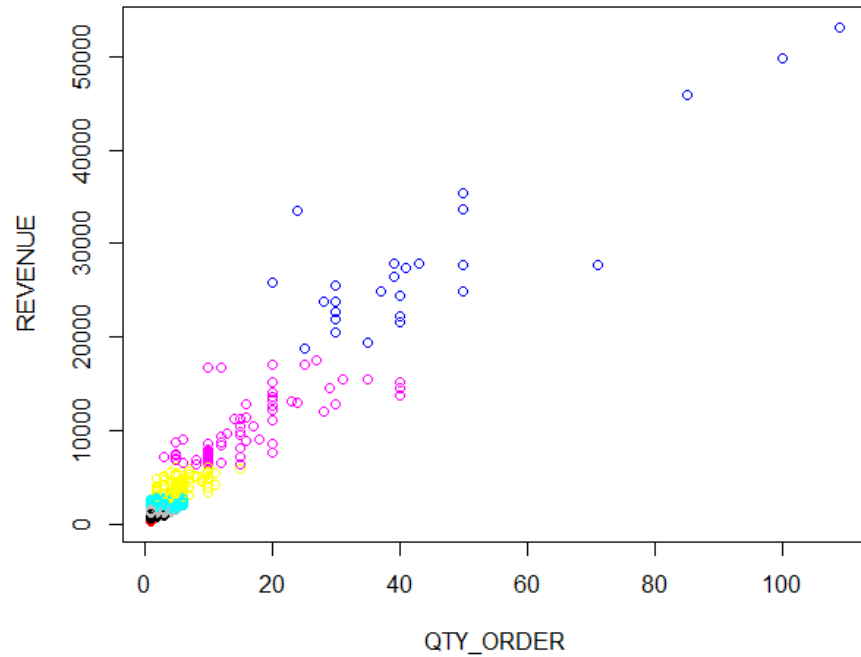
	PROD_ID	DATE_ORDER	QTY_ORDER	REVENUE
<b>P7</b>	:195939	2015-08-19: 12308	Min. : 1.000	Min. : 229
<b>P2</b>	: 63358	2015-08-20: 7737	1st Qu.: 1.000	1st Qu.: 649
<b>P8</b>	: 27840	2015-05-27: 7562	Median : 1.000	Median : 749
<b>P5</b>	: 19775	2015-09-02: 7255	Mean : 1.079	Mean : 765
<b>P9</b>	: 18431	2015-09-01: 6669	3rd Qu.: 1.000	3rd Qu.: 799
<b>P4</b>	: 14656	2015-09-23: 5282	Max. : 500.000	Max. : 368750
(Other):	11092	(Other) : 304278		

## Algumas outras informações...

- Não tinha valor 0 ou faltante nos dados
- A maioria das transações de compra é de um produto
- P7, P8, P9, P5 e P2, tiveram mais transações de compra no mês 8
- P1, P3, tiveram mais transações de compra no mês 3
- P4 e P6, tiveram mais transações de compra mês 9
- P7 teve o maior valor de quantidade vendida em apenas uma transação



## Agrupamento K-Means

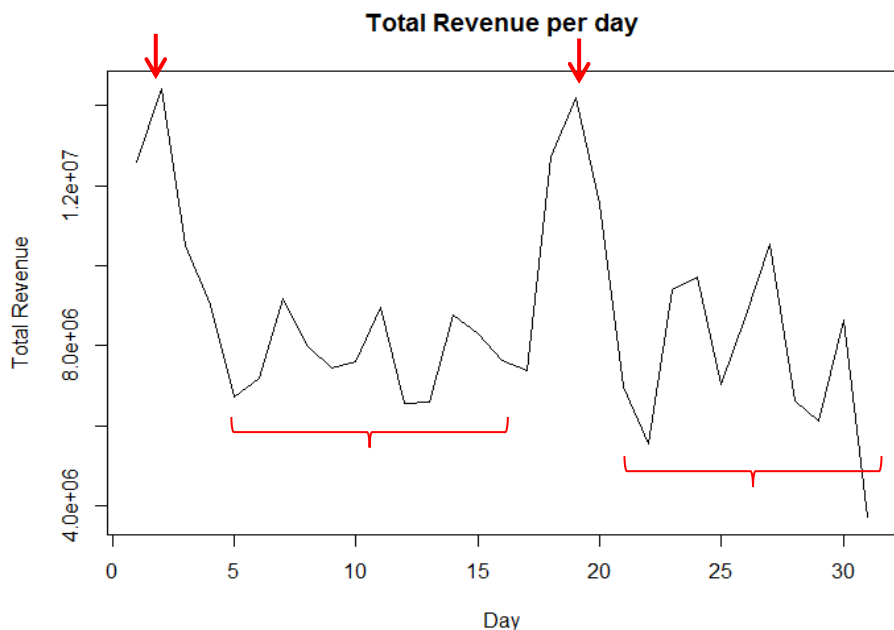
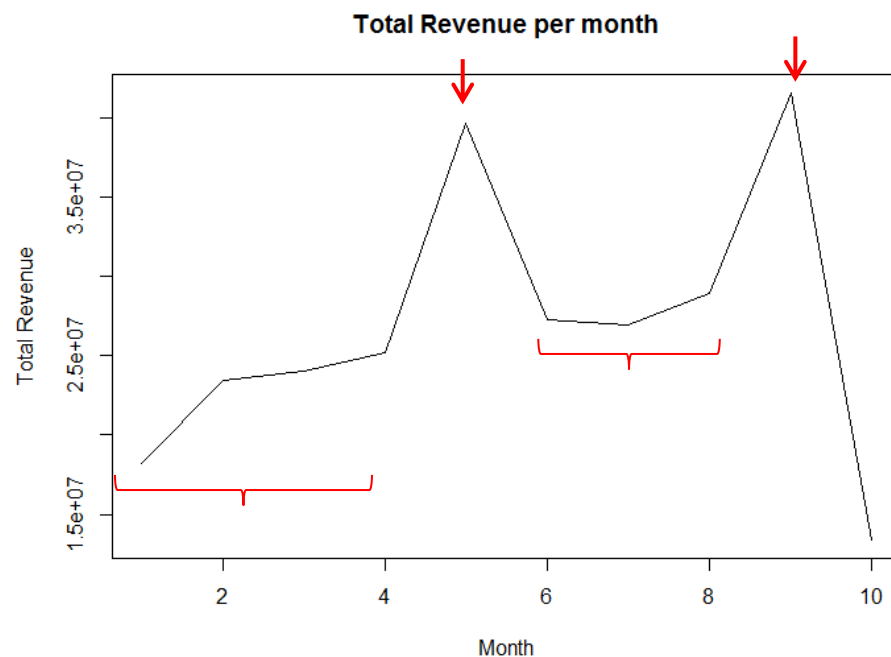


Valor mais alto

Sem o valor mais alto

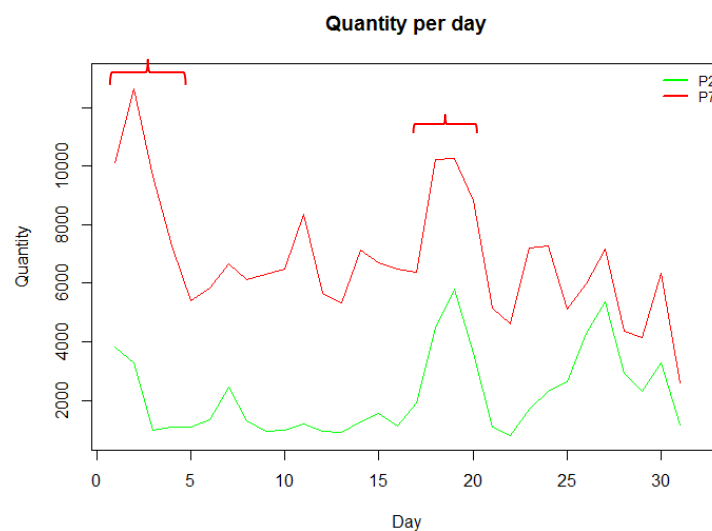
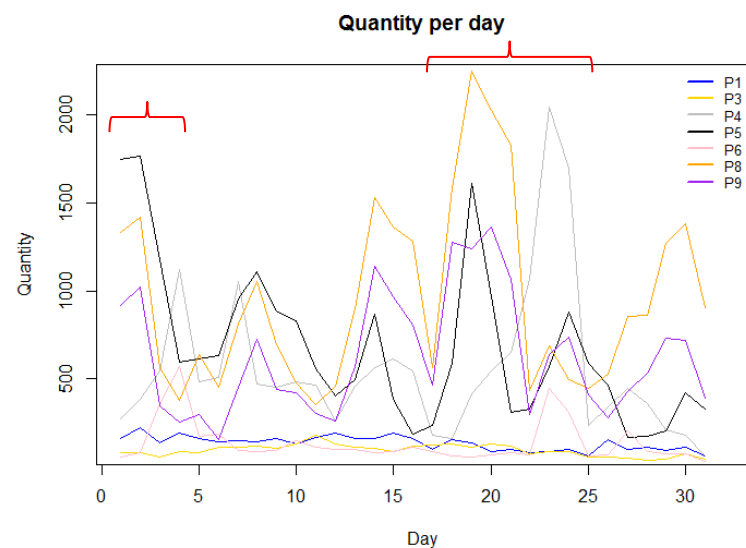
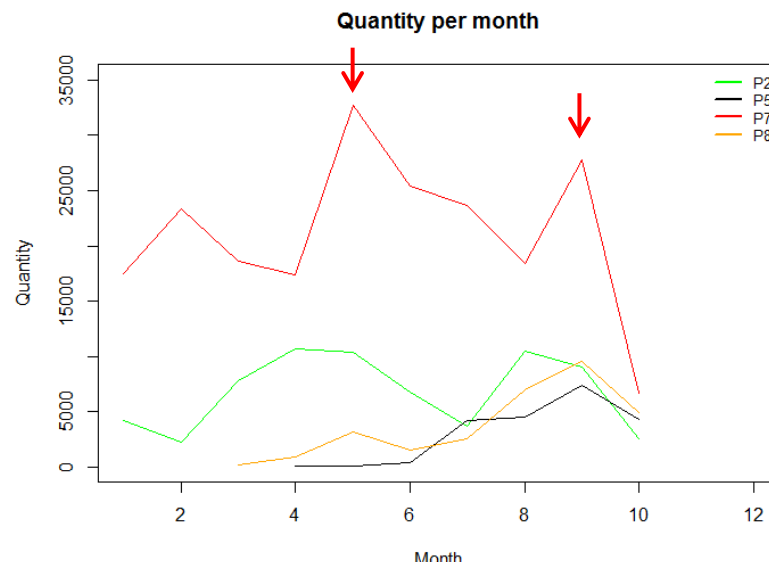
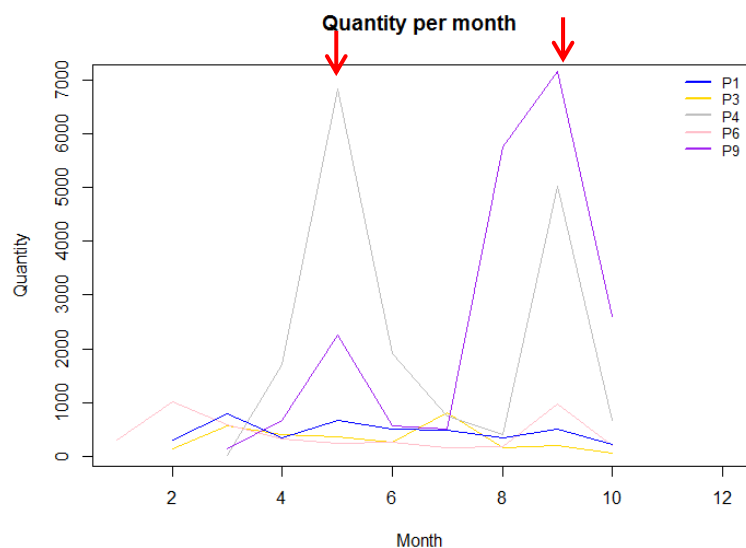
## Visão geral – receita

- Maiores receitas foram obtidas nos meses 5 e 9 (Mai e Set)
- Menores receitas nos meses:
  - 1 a 4 (Jan, Fev, Mar e Abr)
  - 6 a 8 (Jun, Jul e Ago)
  - 10 (Out)



- Maiores receitas foram obtidas nos dias 1 a 3 e 17 a 18
- Menores receitas foram obtidas nos dias 5 a 16 e 21 a 30

## Visão geral – quantidade x mês



## Visão geral – quantidade x dia

Geral:


- Produto que mais foi vendido foi o P7
- Produtos que menos foram vendidos foi o P1, P3 e P6

Mensal:

- As maiores quantidade de produtos vendidos foram nos meses 5 e 9 (Maio e Set)

Diário:

- As maiores quantidade de produtos vendidos foi nos dias 1 a 3 e 17 a 18



Consequentemente  
períodos com  
receitas maiores

Dado um preço, qual a quantidade vendida?

Foi aplicado a Correlação de Pearson entre quantidade e receita (QTY\_ORDER e REVENUE, respectivamente).

O valor obtido foi de **0.948663**.

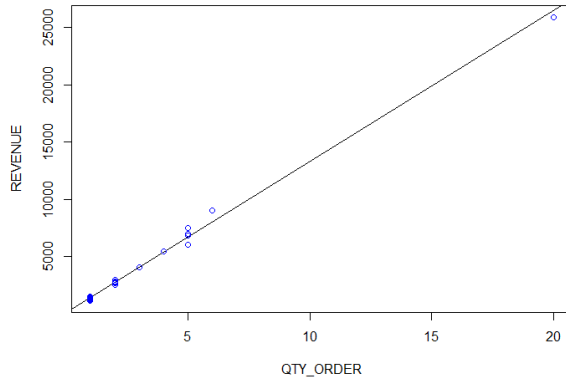
Como esse valor próximo é 1 indica uma forte relação entre as duas variáveis. Como é positivo indica que os valores das duas variáveis tendem a crescer juntas.

Seleção dos dados pelo PROD\_ID

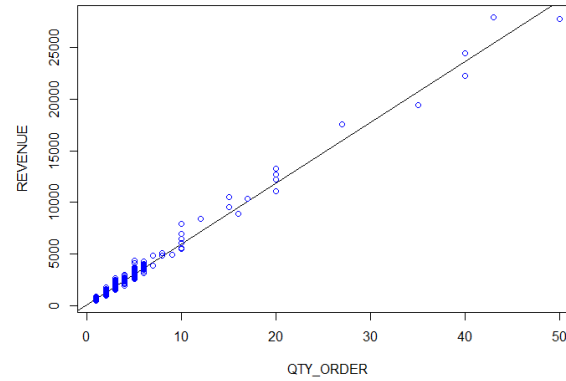
- Aplicação da técnica de regressão linear na relação REVENUE e QTY\_ORDER
- O algoritmo calcula os coeficientes da equação
- Com o modelo criado é submetido o(s) preço(s) e é retornado a(s) quantidade(s) vendida(s).
- Avaliação do modelo criado com a técnica de Validação Cruzada

# 2 – Modelo preditivo

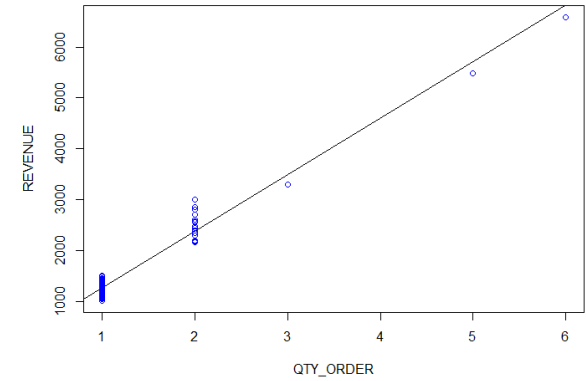
Linear Regression P1



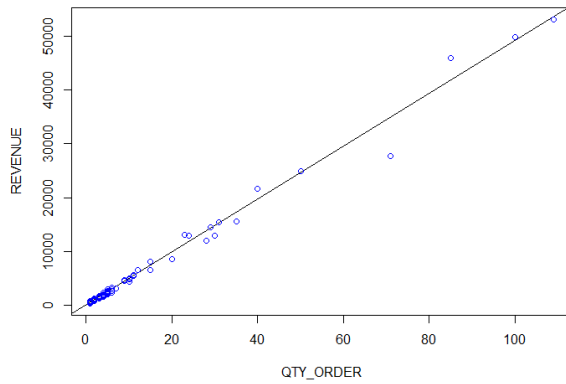
Linear Regression P2



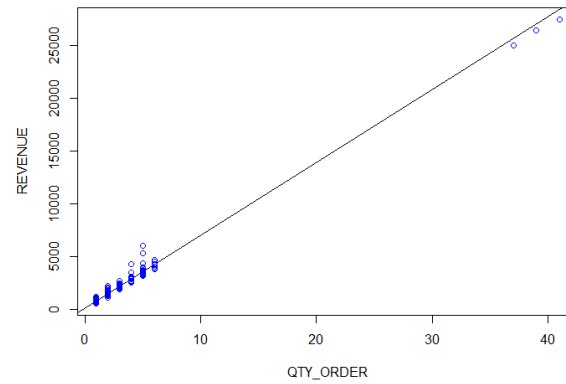
Linear Regression P3



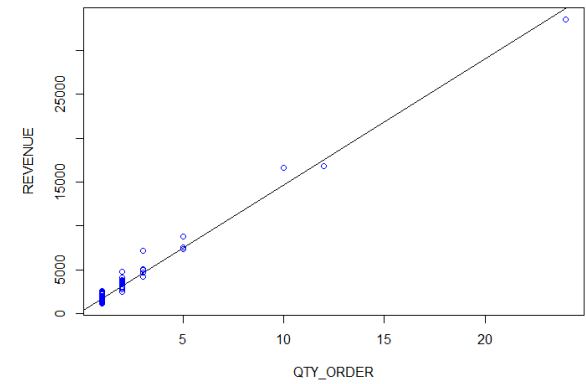
Linear Regression P4



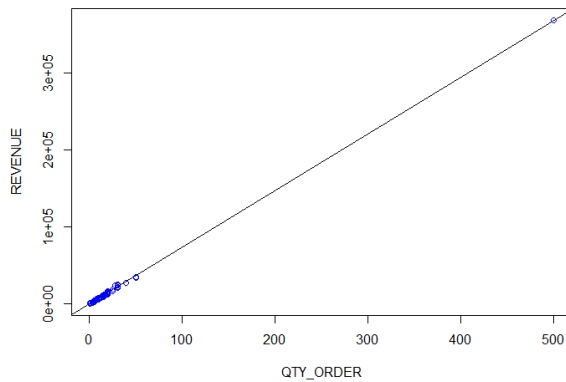
Linear Regression P5



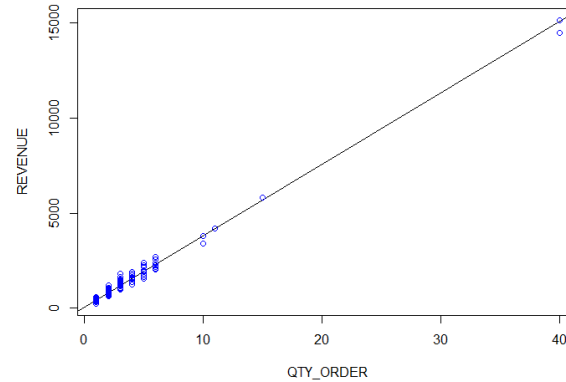
Linear Regression P6



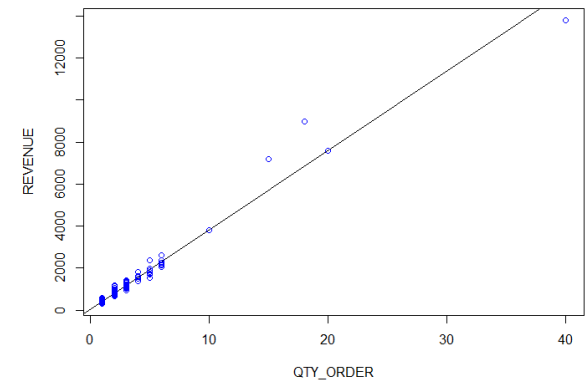
Linear Regression P7



Linear Regression P8



Linear Regression P9



Sumarização: Métricas calculadas para cada produto no conjunto de concorrentes

PROD_ID	DATE_EXTRACTION	COMPETITOR	COMPETITOR_PRICE	PAY_TYPE
P1 :2713	2015-04-24 08:11:36: 12	C1:765	Min. : 1090	Min. :1.000
	2015-07-09 20:52:29: 12	C2:703	1st Qu.: 1424	1st Qu.:1.000
	2015-03-17 09:25:33: 10	C3:396	Median : 1499	Median :1.000
	2015-03-20 08:12:50: 10	C4: 0	Mean : 1908	Mean :1.499
	2015-03-25 08:34:51: 10	C5:713	3rd Qu.: 1499	3rd Qu.:2.000
	2015-04-10 08:11:34: 10	C6:136	Max. :149900	Max. :2.000
	(Other) :2649			
PROD_ID	DATE_EXTRACTION	COMPETITOR	COMPETITOR_PRICE	PAY_TYPE
P2 :8755	2015-01-05 08:11:41: 20	C1:1122	Min. : 506.9	Min. :1.0
	2015-01-05 20:10:17: 20	C2:1128	1st Qu.: 677.9	1st Qu.:1.0
	2015-01-08 08:11:26: 20	C3:2089	Median : 729.5	Median :1.0
	2015-01-08 20:10:13: 20	C4:1959	Mean : 819.9	Mean :1.5
	2015-02-14 08:11:40: 20	C5: 494	3rd Qu.: 799.0	3rd Qu.:2.0
	2015-02-16 20:10:13: 20	C6:1963	Max. :79900.0	Max. :2.0
	(Other) :8635			
PROD_ID	DATE_EXTRACTION	COMPETITOR	COMPETITOR_PRICE	PAY_TYPE
P3 :5853	2015-07-09 20:52:29: 16	C1:1646	Min. : 879.1	Min. :1.000
	2015-04-24 08:11:36: 15	C2:1652	1st Qu.: 1099.0	1st Qu.:1.000
	2015-03-15 20:11:25: 12	C3: 835	Median : 1214.1	Median :1.000
	2015-03-16 08:11:32: 12	C4: 639	Mean : 1421.6	Mean :1.499
	2015-03-16 20:10:14: 12	C5: 286	3rd Qu.: 1312.3	3rd Qu.:2.000
	2015-03-17 09:25:33: 12	C6: 795	Max. :119900.0	Max. :2.000
	(Other) :5774			
PROD_ID	DATE_EXTRACTION	COMPETITOR	COMPETITOR_PRICE	PAY_TYPE
P4 :1689	2015-05-18 20:10:33: 6	C1: 0	Min. : 431.1	Min. :1.0
	2015-05-19 08:11:27: 6	C2: 0	1st Qu.: 497.0	1st Qu.:1.0
	2015-05-19 20:10:34: 6	C3: 0	Median : 499.9	Median :1.0
	2015-05-20 08:11:35: 6	C4:1085	Mean : 643.0	Mean :1.5
	2015-05-20 20:10:27: 6	C5: 16	3rd Qu.: 569.0	3rd Qu.:2.0
	2015-05-21 08:11:51: 6	C6: 588	Max. :49700.0	Max. :2.0
	(Other) :1653			

### 3 – Análise concorrência e vendas

PROD_ID	DATE_EXTRACTION	COMPETITOR	COMPETITOR_PRICE	PAY_TYPE
P5 :1896	2015-07-09 20:52:29: 12	C1:628	Min. : 674.1	Min. :1.0
	2015-05-04 20:10:42: 6	C2:636	1st Qu.: 809.1	1st Qu.:1.0
	2015-05-05 08:11:53: 6	C3:632	Median : 886.5	Median :1.5
	2015-05-05 20:10:25: 6	C4: 0	Mean : 1142.3	Mean :1.5
	2015-05-06 08:11:29: 6	C5: 0	3rd Qu.: 933.2	3rd Qu.:2.0
	2015-05-07 08:12:29: 6	C6: 0	Max. :84890.0	Max. :2.0
(Other) :1854				

PROD_ID	DATE_EXTRACTION	COMPETITOR	COMPETITOR_PRICE	PAY_TYPE
P6 :9542	2015-01-08 08:11:26: 20	C1:2210	Min. : 1226	Min. :1.0
	2015-01-08 20:10:13: 20	C2:2142	1st Qu.: 1674	1st Qu.:1.0
	2015-02-14 08:11:40: 20	C3:2166	Median : 1799	Median :1.5
	2015-02-16 20:10:13: 20	C4:1190	Mean : 1952	Mean :1.5
	2015-02-17 08:10:23: 20	C5: 0	3rd Qu.: 1928	3rd Qu.:2.0
	2015-02-17 08:11:34: 20	C6:1834	Max. :149900	Max. :2.0
(Other) :9422				

PROD_ID	DATE_EXTRACTION	COMPETITOR	COMPETITOR_PRICE	PAY_TYPE
P7 :7748	2015-02-14 08:11:40: 24	C1: 991	Min. : 588.7	Min. :1.0
	2015-01-08 20:10:13: 20	C2: 979	1st Qu.: 745.0	1st Qu.:1.0
	2015-02-16 20:10:13: 20	C3:1077	Median : 788.0	Median :1.0
	2015-03-05 20:11:17: 20	C4:2249	Mean : 893.5	Mean :1.5
	2015-04-24 08:11:36: 20	C5:1549	3rd Qu.: 849.0	3rd Qu.:2.0
	2015-07-02 08:11:22: 18	C6: 903	Max. :104900.0	Max. :2.0
(Other) :7626				

PROD_ID	DATE_EXTRACTION	COMPETITOR	COMPETITOR_PRICE	PAY_TYPE
P8 :5795	2015-07-09 20:52:29: 16	C1:1253	Min. : 359.1	Min. :1.0
	2015-09-21 19:06:18: 16	C2:1263	1st Qu.: 431.1	1st Qu.:1.0
	2015-09-02 20:10:08: 15	C3:1253	Median : 448.0	Median :1.0
	2015-05-05 08:11:53: 12	C4: 863	Mean : 509.5	Mean :1.5
	2015-05-05 20:10:25: 12	C5: 14	3rd Qu.: 479.0	3rd Qu.:2.0
	2015-05-06 08:11:29: 12	C6:1149	Max. :39999.0	Max. :2.0
(Other) :5712				

PROD_ID	DATE_EXTRACTION	COMPETITOR	COMPETITOR_PRICE	PAY_TYPE
P9 :6123	2015-09-21 19:06:18: 20	C1:1253	Min. : 359.1	Min. :1.0
	2015-07-09 20:52:29: 16	C2:1267	1st Qu.: 431.1	1st Qu.:1.0
	2015-09-02 20:10:08: 15	C3:1247	Median : 448.2	Median :1.0
	2015-05-14 20:10:06: 12	C4: 4	Mean : 533.7	Mean :1.5
	2015-05-15 08:10:23: 12	C5:1215	3rd Qu.: 496.0	3rd Qu.:2.0
	2015-05-04 20:10:42: 10	C6:1137	Max. :56900.0	Max. :2.0
(Other) :6038				



### 3 – Análise concorrência e vendas

#### Considerações

Valor suspeito

PROD_ID	DATE_EXTRACTION	COMPETITOR	COMPETITOR_PRICE	PAY_TYPE	max(COMPETITOR_PRICE)	
1	P1	2015-10-14 08:11:39	C2	149900	2	149900
PROD_ID	DATE_EXTRACTION	COMPETITOR	COMPETITOR_PRICE	PAY_TYPE	max(COMPETITOR_PRICE)	
1	P2	2015-10-14 08:11:39	C4	79900	2	79900
PROD_ID	DATE_EXTRACTION	COMPETITOR	COMPETITOR_PRICE	PAY_TYPE	max(COMPETITOR_PRICE)	
1	P3	2015-10-14 08:11:39	C2	119900	2	119900
PROD_ID	DATE_EXTRACTION	COMPETITOR	COMPETITOR_PRICE	PAY_TYPE	max(COMPETITOR_PRICE)	
1	P4	2015-10-14 08:11:39	C6	49700	2	49700
PROD_ID	DATE_EXTRACTION	COMPETITOR	COMPETITOR_PRICE	PAY_TYPE	max(COMPETITOR_PRICE)	
1	P5	2015-10-14 08:11:39	C2	84890	1	84890
PROD_ID	DATE_EXTRACTION	COMPETITOR	COMPETITOR_PRICE	PAY_TYPE	max(COMPETITOR_PRICE)	
1	P6	2015-10-14 08:11:39	C6	149900	1	149900
PROD_ID	DATE_EXTRACTION	COMPETITOR	COMPETITOR_PRICE	PAY_TYPE	max(COMPETITOR_PRICE)	
1	P7	2015-10-14 08:11:39	C3	104900	2	104900
PROD_ID	DATE_EXTRACTION	COMPETITOR	COMPETITOR_PRICE	PAY_TYPE	max(COMPETITOR_PRICE)	
1	P8	2015-10-14 08:11:39	C6	39999	2	39999
PROD_ID	DATE_EXTRACTION	COMPETITOR	COMPETITOR_PRICE	PAY_TYPE	max(COMPETITOR_PRICE)	
1	P9	2015-10-14 08:11:39	C5	56900	1	56900

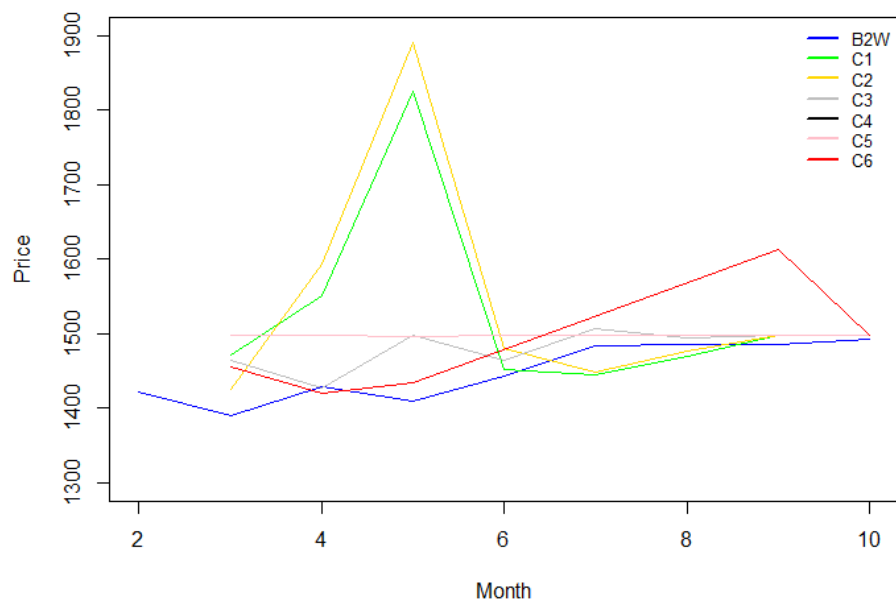
Valores  
retirados da  
análise

Relação de preços {  
Comp\_Prices: média do preço do produto  
Sales: preço=receita/quantidade

Consideração sobre concorrente é mais importante → tendência parecida no gráfico, proximidade ou menor preço para cada produto

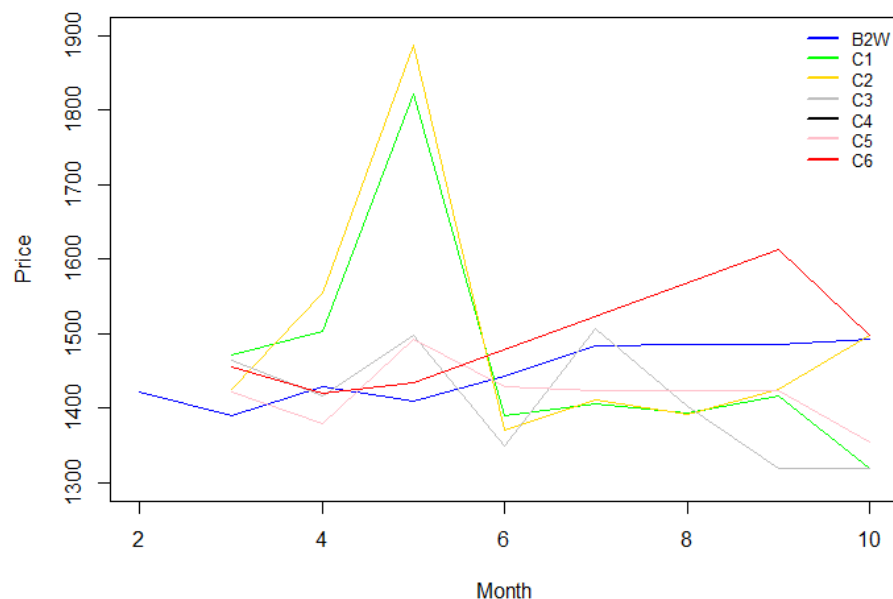
## P1

Product P1 and Pay type 1



- C3 com valores parecidos

Product P1 and Pay type 2



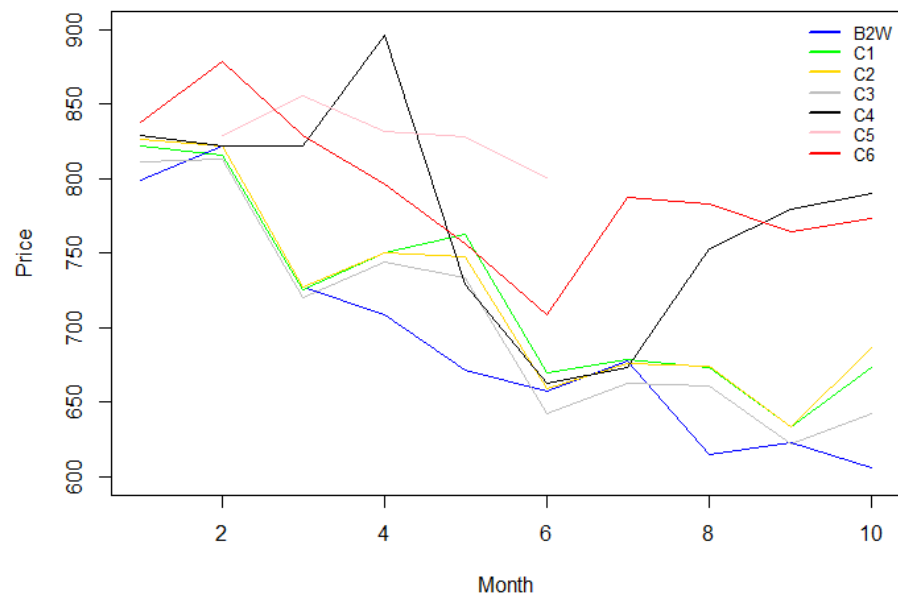
- C3 e C5 com valores próximos ou menores

- Ausência de C4
- Início de monitoramento em março

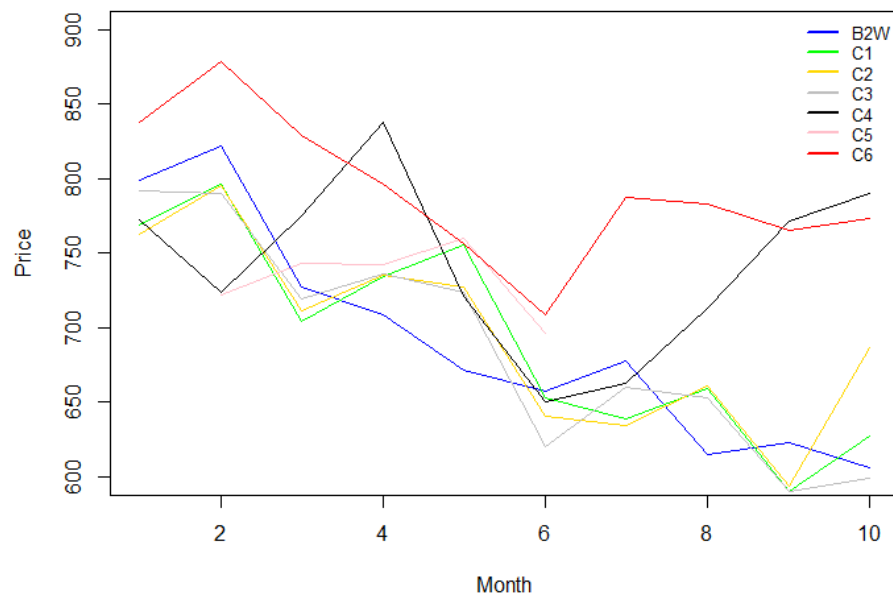
Apesar de variar ao longo dos meses, pode-se considerar como concorrente mais importante para venda de P1 é C3

## P2

Product P2 and Pay type 1



Product P2 and Pay type 2

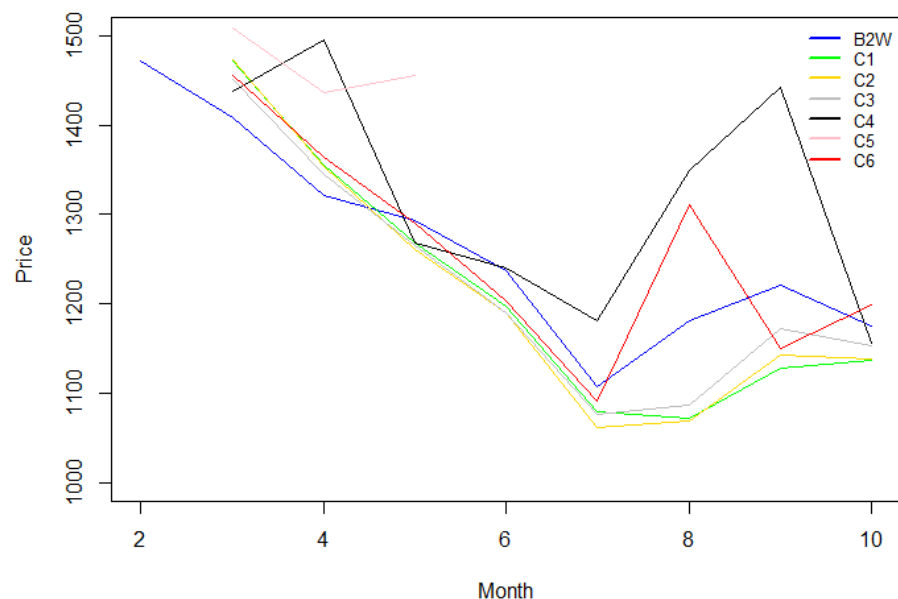


- C3 com preço menor em Jun e Jul
- C1, C2 e C3 com preço menor que B2W em vários períodos
- C5 com dados em Fev a Jun, somente
- C1, C2 e C3 com valores parecidos ou menores que B2W

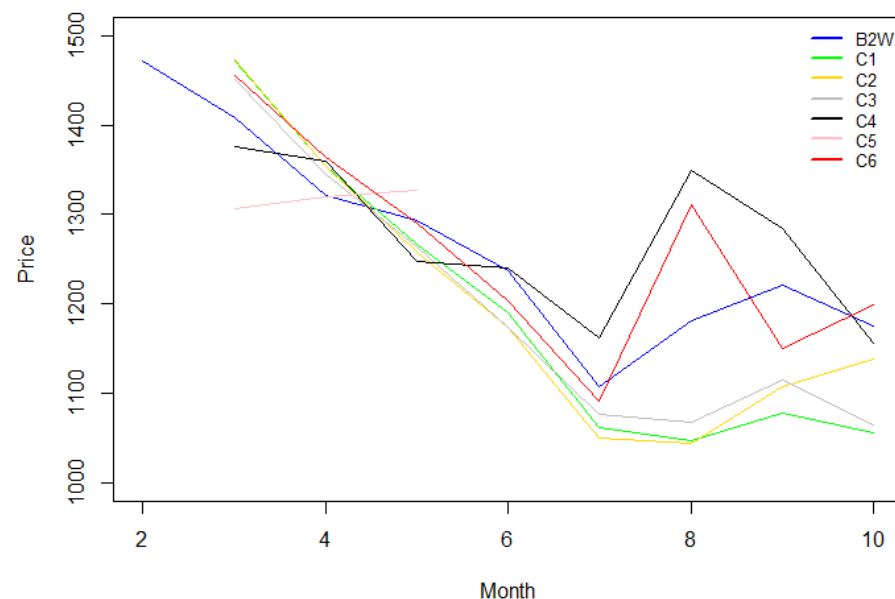
Pode-se considerar como concorrente mais importante para venda de P2 é **C3**

## P3

Product P3 and Pay type 1



Product P3 and Pay type 2

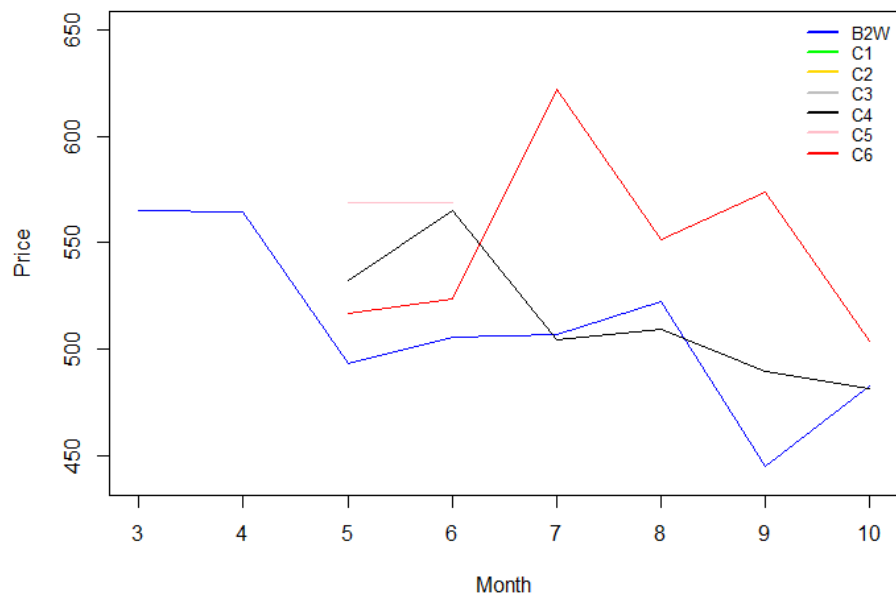


- Dados do monitoramento a partir de Mar
- C5 com dados em Mar, Abr e Mai, somente
- C1, C2, C3 e C6 com valores próximos ou menores que B2W

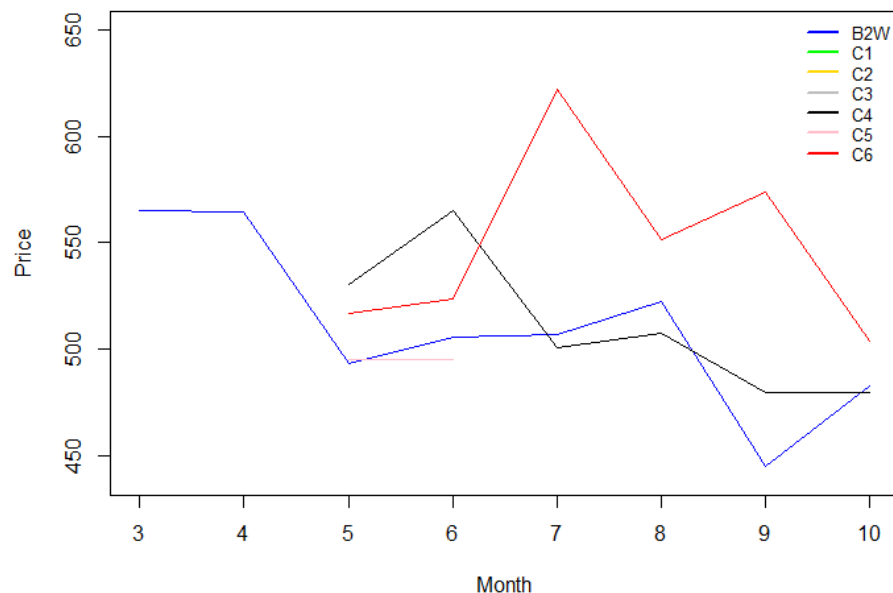
Pode-se considerar como concorrente mais importante para venda de P3 são **C1, C2 e C3**

## P4

Product P4 and Pay type 1



Product P4 and Pay type 2

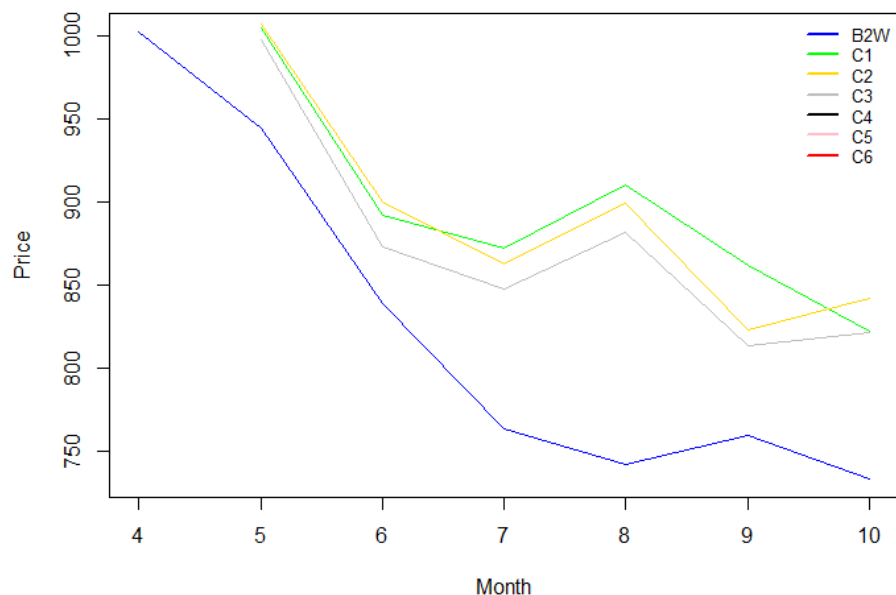


- C4 apresentou preços menores que B2W em Jul e Ago
  - Início de monitoramento em Mai
    - Ausência de C1, C2 e C3
    - C5 presente somente em Mai
  - Tipo de pagamento 1 e 2 parecidos
- C4 e C5 tiveram preços menores em alguns períodos

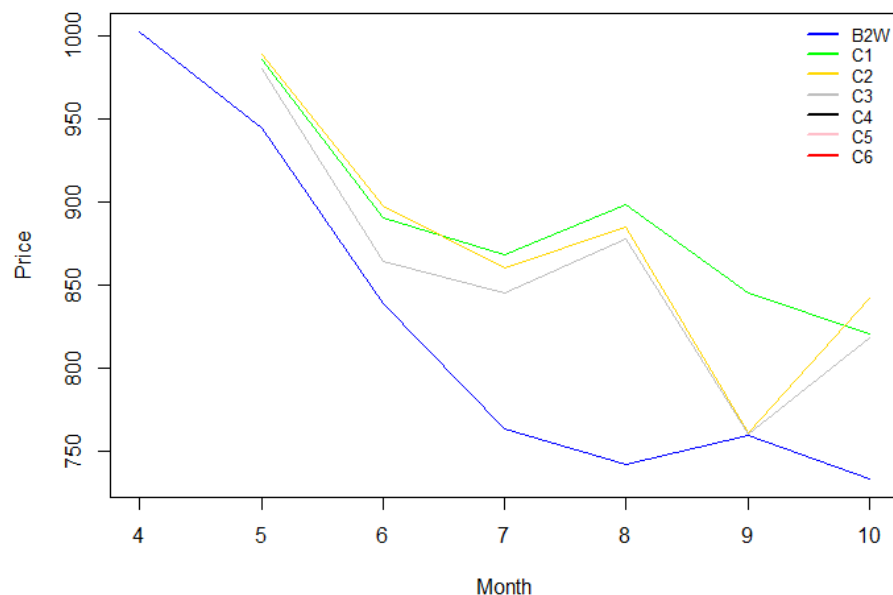
Apesar de variar ao longo dos meses, pode-se considerar como concorrente mais importante para venda de P4 é C4

## P5

Product P5 and Pay type 1



Product P5 and Pay type 2

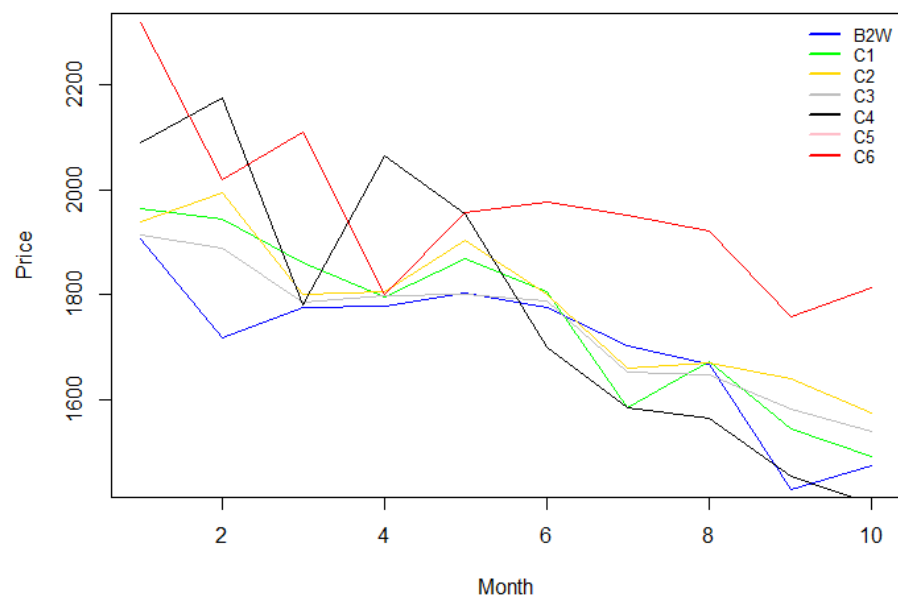


- Início de monitoramento em Mai
  - Ausência de C4, C5 e C6
  - B2W com preço menor em todo o período de análise
- C2 e C3 tiveram o mesmo preço que B2W somente em setembro (tipo 2)

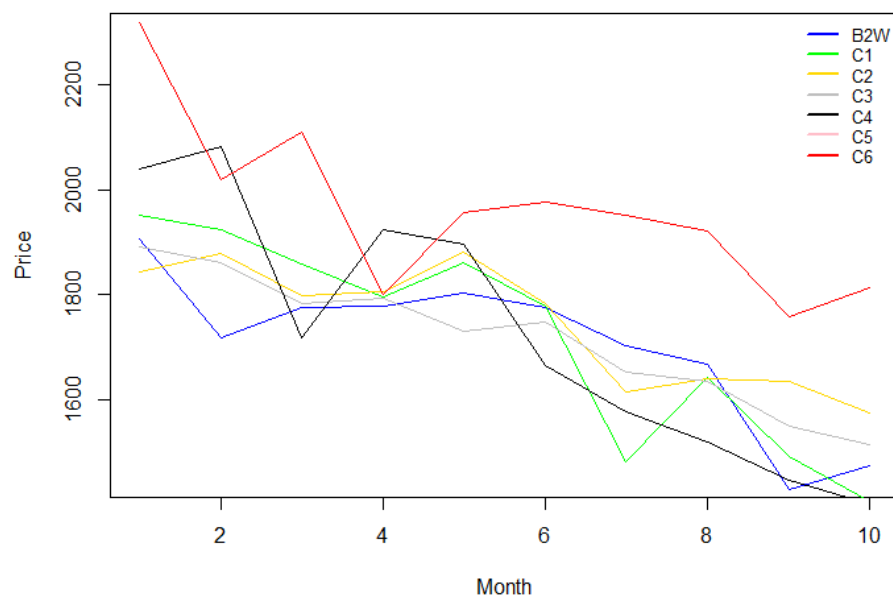
Pode-se considerar como concorrente mais importante para venda de P5 é C3, devido a sua maior proximidade de valores

## P6

Product P6 and Pay type 1



Product P6 and Pay type 2

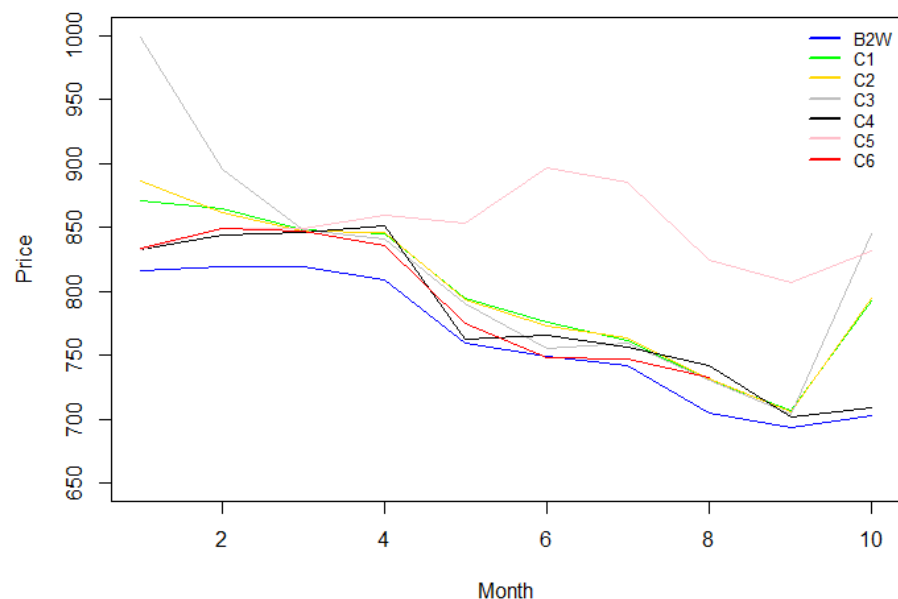


- Ausência de C5
- C1, C2, C3 e C4 com valores parecidos e/ou menores que B2W em alguns períodos

Apesar da variação entre os meses, pode-se considerar como concorrente mais importante para venda de P6 são C1, C2, C3 e C4.

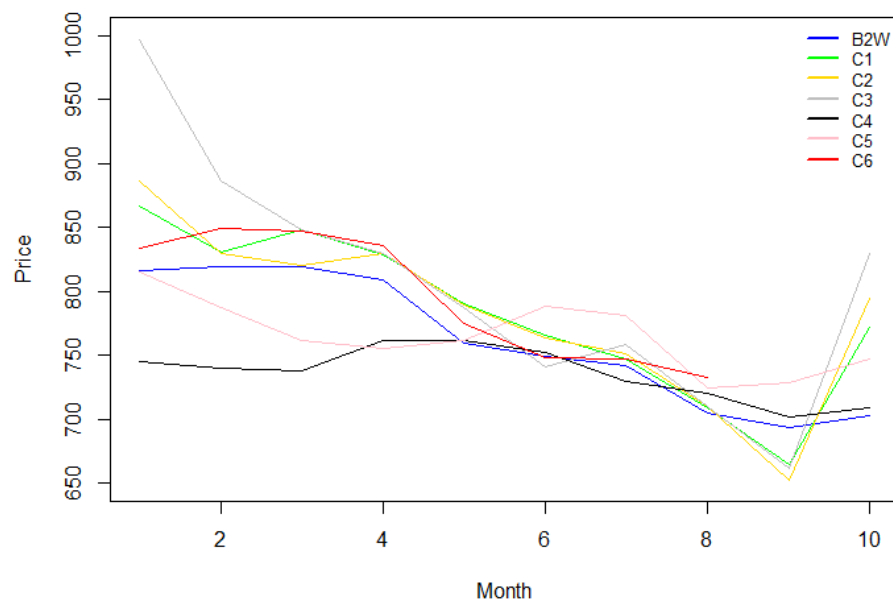
## P7

Product P7 and Pay type 1



- B2W apresentou preços menores
- C4 teve valores próximos

Product P7 and Pay type 2



- C5 e C4 tiveram preços menores de Jan a Jun
- C2 apresentou valores próximos

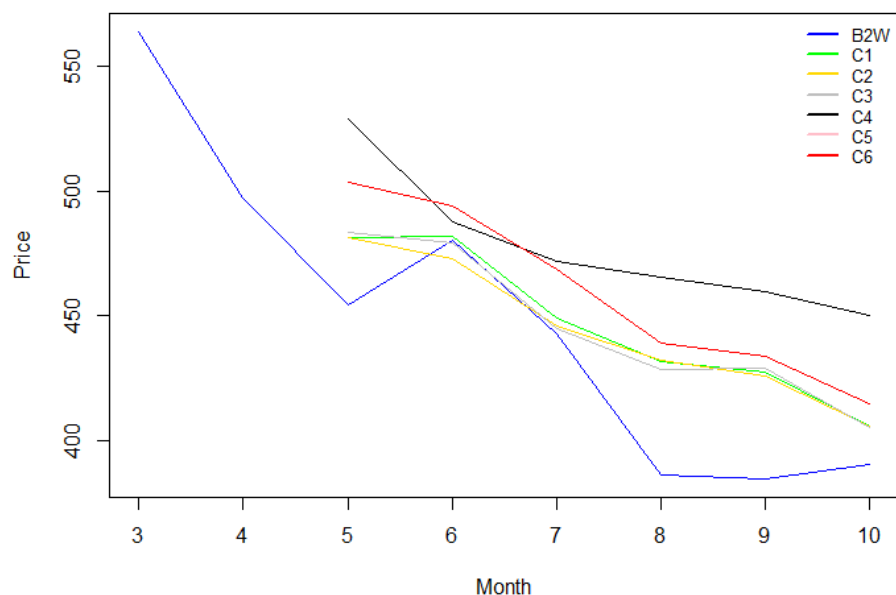
- C6 teve valores próximos, mas foi até Ago

Em P7, pode-se considerar como concorrente mais importante o **C4**.

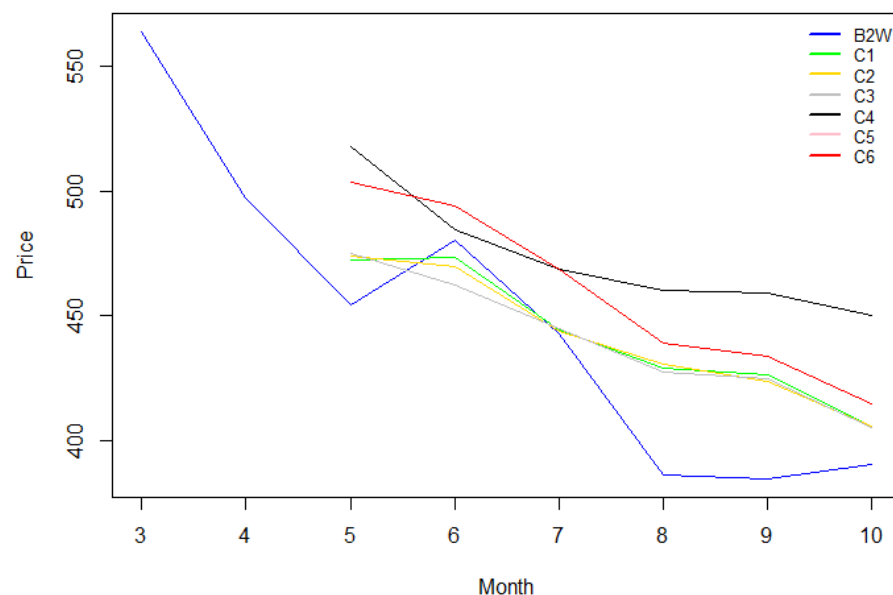


## P8

Product P8 and Pay type 1



Product P8 and Pay type 2

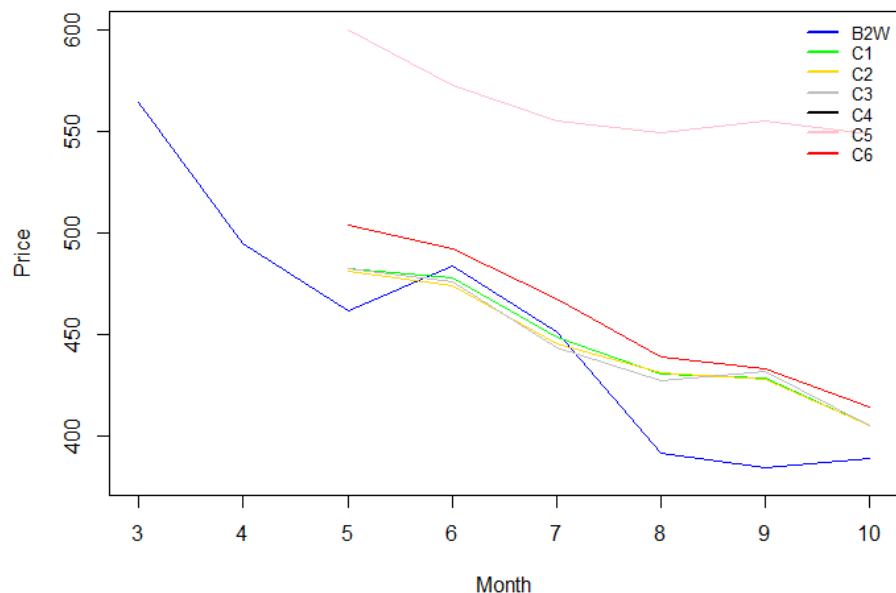


- Monitoramento iniciado no mês 5
- De Mai a Ago os competidores C1, C2 e C3 tiveram valores mais próximos da B2W
  - Tipo de pagamento 1 e 2 parecidos

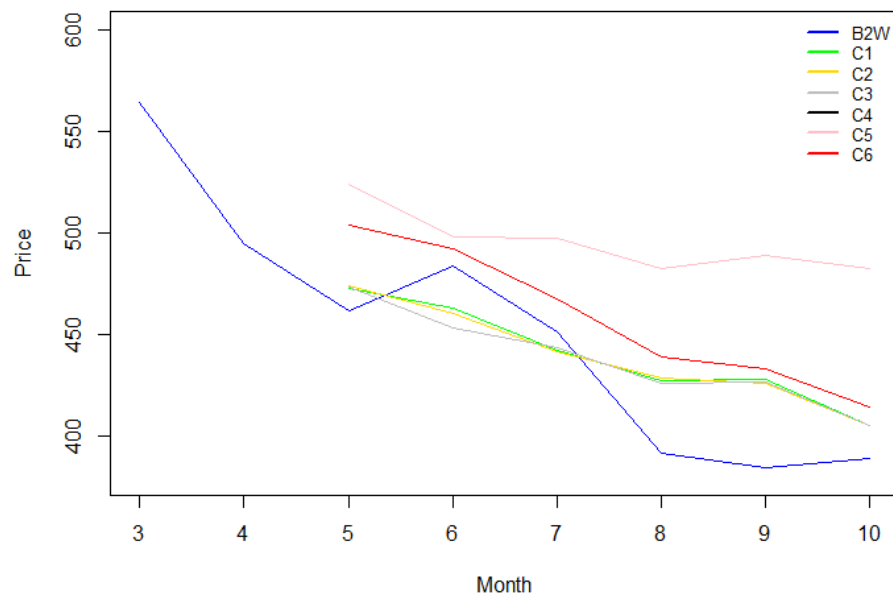
Em P8, pode-se considerar como concorrentes mais importante o **C1**, **C2** e **C3**.

## P9

Product P9 and Pay type 1



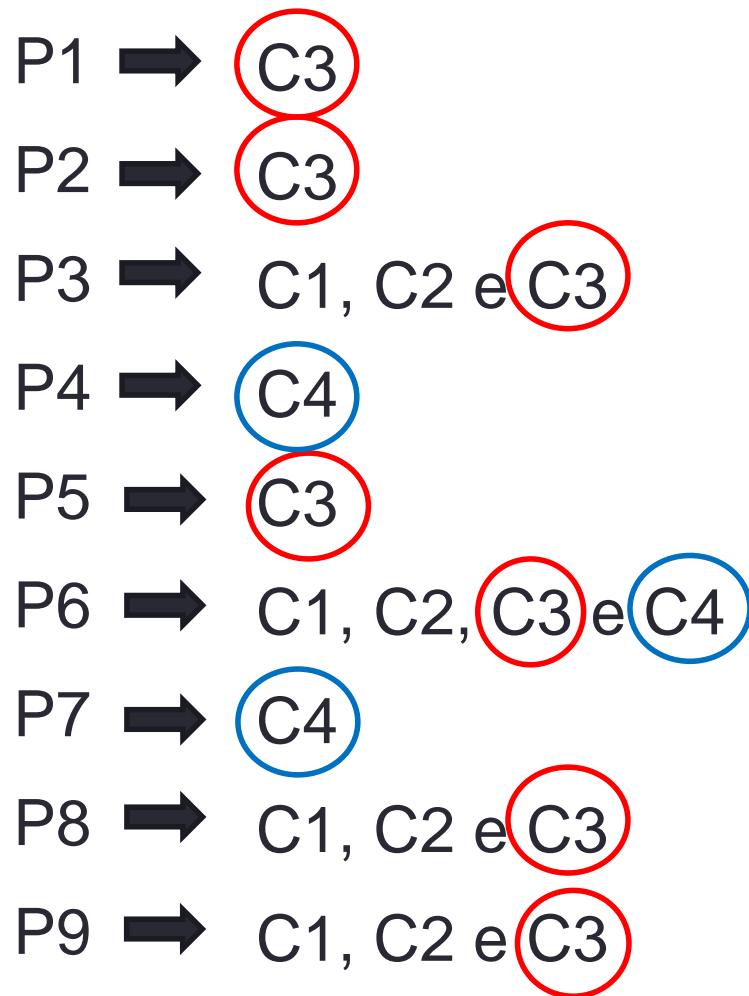
Product P9 and Pay type 2



- Monitoramento iniciado no mês 5
- No início, de Mai a Ago os competidores C1, C2 e C3 tiveram valores mais próximos da B2W

Em P8, pode-se considerar como concorrentes mais importante o **C1, C2 e C3**

#### Geral



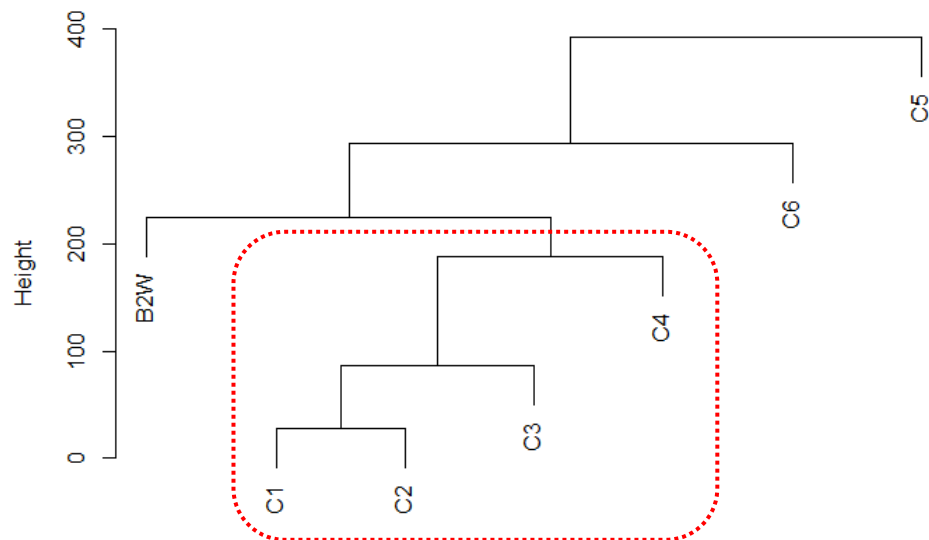
Destaque:

- Ausência de C6 e C5
- P4 e P7 o concorrente foi o C4
- No restante do produtos foi o competidor C3

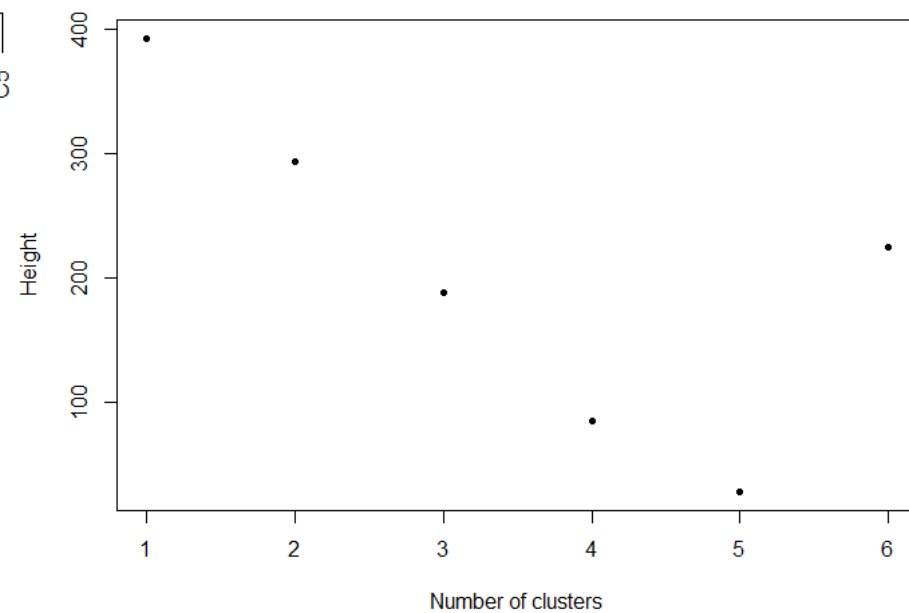
Concorrente mais importante: aparecem mais vezes que aparecem de maneira única em um produto.

## Agrupamento - Agnes

Dendrogram of `agnes(x = data_full, method = "average")`



Grupo de concorrentes  
mais próximo



**Obrigado pela atenção.**