

## 2. A relación entre prezo e peso.

### 2.1 Relación simple.

En todas as situacións analizadas no apartado anterior vese como as diferencias no prezo dos diamantes en diferentes situacións parece facilmente xustificable polas diferencias de peso, por que soen aparecer asociadas (*táboas 3 e 4*).

Para comprobar isto vaise facer unha análise bidimensional, na que se estudará a relación entre esas dúas variables, observando tamén como lle afectan as diferencias nas propiedades consideradas.

Esta suposta relación pode verse na *gráfica 5*, xa que os puntos definen bastante ben unha liña recta. Tamén se observan outras cousas, como unhas maiores diferencias entre os prezos para pesos maiores e dous rangos de pesos nos que apenas hai diamantes, o que explicaría os 3 máximos que observamos ao representar o diagrama do peso (*gráfica 4*).

Nesa gráfica tamén se observa en que zonas a relación entre prezo e peso se aparta desa liña recta: ao inicio, con pesos e prezos baixos, onde o incremento do prezo en función do peso é máis suave, e ao final, onde existe moitas diferencias no prezo para pesos semellantes.

A regresión para estas variables da o seguinte resultado:

$$\text{Prezo} = -2274 + 12621 * \text{Peso}$$

Cada unidade de peso aumenta o prezo en 12 621 \$, aínda que quizais sería mellor

---

expresalo en décimas, posto que 1 unidade de peso case abrangue todo o rango dos datos; podería dicirse polo tanto que por cada décima máis de peso agárdase un aumento de 1261,2 \$. O término independente negativo é consecuencia da diferenza na relación nos prezos baixos, onde parece que o incremento se produce máis suavemente.

Esta regresión é moi boa posto que  $r^2=0,877$ , ou sexa, conseguimos explicar coa regresión o 87,7% da variación nos prezos.

## 2.2 Influencia das propiedades na relación.

Para intentar mellorar o resultado obtido co modelo simple calculouse a mesma regresión pero separando os datos en función dos diferentes niveis das propiedades.

O resume dos resultados pode verse na *táboa 7*. A mellora é evidente posto que se mellora o valor do  $r^2$ , subindo como mínimo a variación explicada a un 94%.

Observando os coeficientes vese que o incremento do prezo e función do peso ten un máximo para a cor D, con todo, debemos descartar este resultado posto que só dispoñemos de 3 diamantes con esa cor. Para as demais cores o incremento separaríase en 2 grupos: as cores E e F que teñen un maior incremento e os demais que aínda mantendo un coeficiente próximo ao global, non medra con tanta rapidez pois pasa dos 12 261 globais a cantidades de en torno aos 10 000\$ por cada unidade de peso.

Con todo, esta mellora na explicación ten un matiz de dúbida, posto que analizando a *gráfica 8*, pode verse que a existencia de mais ou menos puntos na zona de pesos máximos está relacionada coas pendentes anteriores: as cores con maior pendente amosan un comportamento similar aos demais, pero teñen algúns puntos na zona de pesos máximos que poden ser a causa da súa diferenza, e non a existencia de diferentes valoracións por cores, que era o que pretendíamos confirmar.

Con respecto a relación coa claridade (*táboa 6*) tamén se mellor a explicación, pero sen que as diferenzas observadas sexan tan grandes coma no caso das cores.

Un intento máis de explicar o noso modelo fíxose transformando o atributo CLARIDADE nunha variable numérica (asignando valor 0 a IF e 1 a VVS1) e calculando un modelo con regresión múltiple. O resume dos resultados está na *táboa 7*, e nel vese que a mellora conseguida é insignificante con respecto ao modelo simple de prezo en función do peso, polo que se descartou o seu uso.

## Anexos: Táboas

PREZO			
COR	a	b	r <sup>2</sup>
D	-8469	24146	0,999
E	-2259	14269	0,951
F	-2391	13487	0,947
G	-1519	10654	0,973
H	-1657	10008	0,951
I	-2000	10635	0,984

Táboa 5

PESO			
Claridade	a	b	r <sup>2</sup>
D	-1933	12447	0,956
E	-3284	13835	0,915

Táboa 6

PESO			
	b <sub>0</sub>	Peso	Claridade
Prezo	-2225,4	13232,8	-673,4

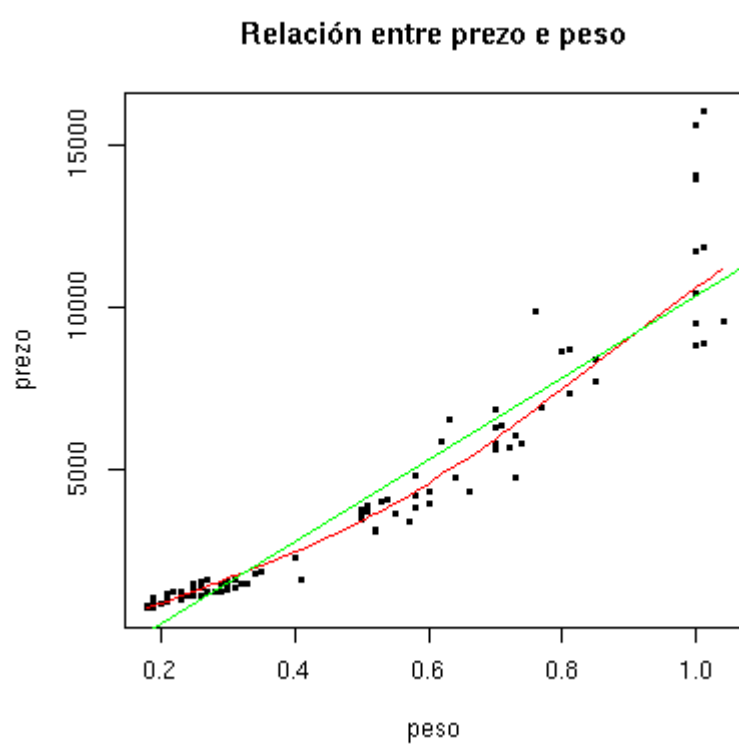
$$R^2 = 0,884$$

$$R^2 \text{ axustado} = 0,881$$

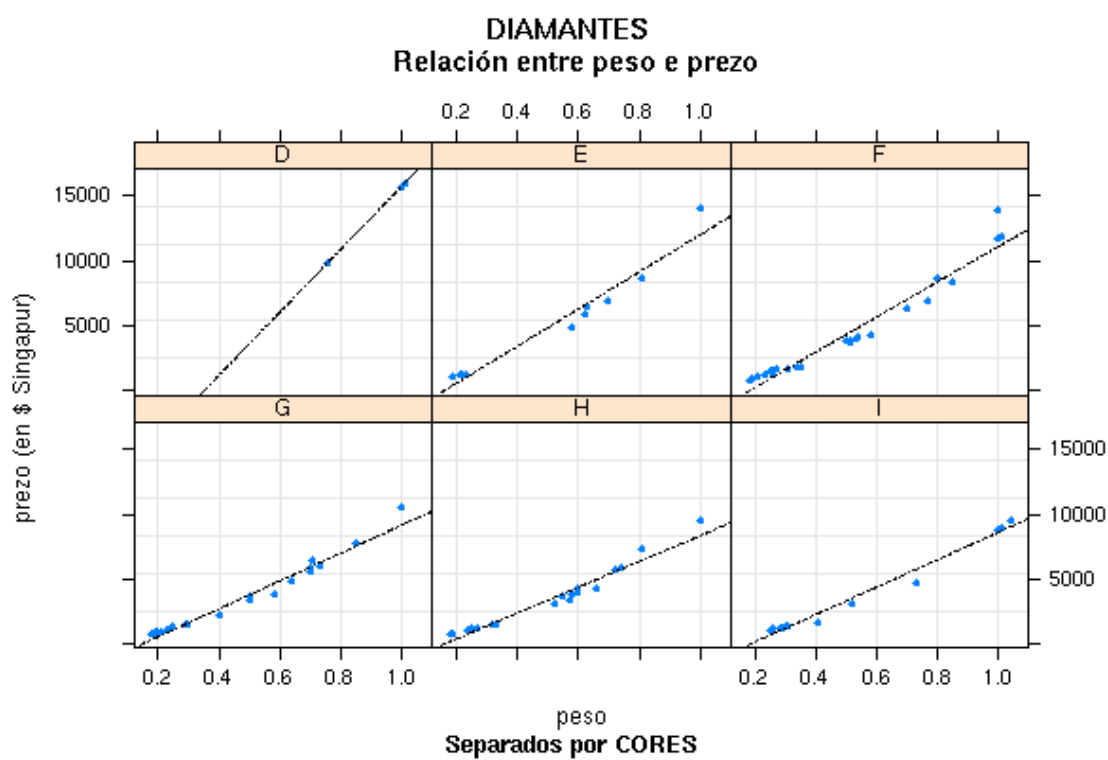
Táboa 7

---

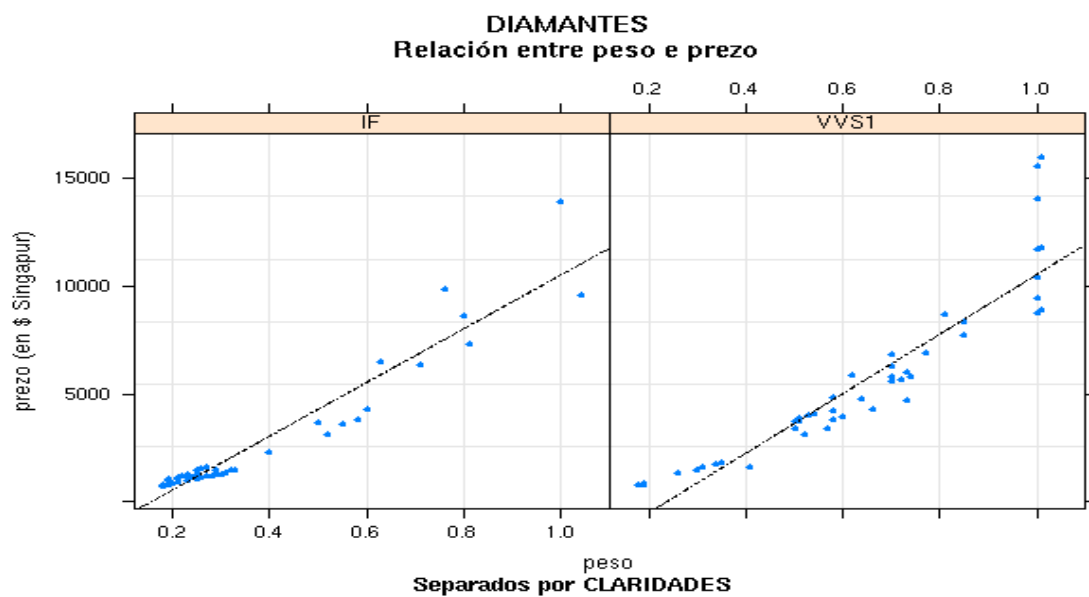
## Anexos: Gráficas



Gráfica 5



*Gráfica 6*



*Gráfica 7*